ASIA PACIFIC JOURNAL
OF EDUCATORS AND EDUCATION

Volume 40, Issue 3, 2025
https://doi.org/10.21315/apjee2025.40.3.11

Research Article:

EdQUAL – A Paradigm to Measure Overall Quality in Education: Empirical Evidence from Vietnamese Lower Secondary Schools

Anh-Duc Hoang[1]* Thanh-Thao Phan[2], Yen-Chi Nguyen[1,3], Viet-Anh Duong Phu[4] and Ngoc-Thuy Ta[5]

[1]RMIT University, 702 Nguyen Van Linh, District 7, Ho Chi Minh City, Viet Nam
[2]Thanh Do University, Km15, National Highway 32, Hoai Duc, Hanoi, Viet Nam
[3]Hanoi University, 51 Đuong Phung Khoang Hanoi 12011 Hanoi, Viet Nam
[4]Independent researcher, Hanoi, Viet Nam
[5]EdLab Asia Educational Research and Development Centre, 83, Nguyen Khang Street, Yen Hoa Ward, Hanoi City, Hanoi, Vietnam

*Corresponding author: duc.hoang8@rmit.edu.vn

ABSTRACT

"Quality in education" is an ambitious objective that every education entity or system wants to pursue, regardless of levels of governance, legislation, diversity of motivation, and approaches towards education. However, the alignment on how we should define quality in education is not reached yet. That challenge in illustrating a portrait of quality education also makes it more difficult for school leaders and education system leaders to improve education quality. For hundreds of years, education systems worldwide witnessed the blooming of theories and practices to tackle this Gordian knot. To be dated, most of the recent approaches focus on the macro perspective of an education system and demand high-level technical skills to implement. Whilst, there are limited well-known guidelines and practices for school leaders to capture the overall picture of contemporary quality issues at their schools for instant and continuous improvements. As a small puzzle to contribute to the fourth United Nations Sustainable Development Goal (SDG 4), this research proposes a framework to measure overall quality in education using a lean and rounded perspective, which can be adopted by any school, school district, or education system. Besides, using the data of 2,239 students' expectations and perceptions about the quality in education, this article also presents the empirical evidence of eight Vietnamese lower secondary schools as a case study.

Keywords: Education quality, SDG 4, EdQUAL, education management, total quality management

Accepted: 14 July 2025; **Published:** 31 December 2025

To cite this article: Hoang, A. D., Phan, T. T., Nguyen, Y. C., Duong Phu, V. A., & Ta, N. T. (2025). EdQUAL – A paradigm to measure overall quality in education: Empirical evidence from Vietnamese lower secondary schools. *Asia Pacific Journal of Educators and Education, 40*(3), 273–305. https://doi.org/10.21315/apjee2025.40.3.11

INTRODUCTION

Sustainable education is an essential target of any nation all over the world (Stabback, 2016). Thus, measuring the quality of an education program is inevitable (Hoang et al., 2020). Regardless of the formal or non-formal context of a school or an out-of-school academic hub, education quality measurement always serves as a focal point to all educational stakeholders (Reimers & Tiburcio, 1993). Among those stakeholders, school leaders and school managers can be benefited greatly from the quality measurement tool. It can produce a lot of information and subsequently will help the leaders to understand, evaluate, and control the school's quality (Craig, 2021; Smith, 1996). Even though nowadays school leaders have to deal with many more problems (Armstrong et al., 2021), "measuring quality" is still a stubborn puzzle that educational experts have been trying to solve (Yocke, 1997). Therefore, this research aims to establish a measurement framework for the overall quality in education of K-12 education institution(s) as well as report the related empirical evidence of Vietnamese lower secondary schools.

In a regular education system, teacher effectiveness, student performance, and school improvement are the top overriding factors whenever mentioning the quality in education measurement (Leu & Price-Rom, 2006). For example, in a three-dimensional relationship, school quality was raised when schools created more learning opportunities for teachers through education or training coursework, which simultaneously improved teacher's subject matter knowledge and student learning results (Guyton & Farokhi, 1987). In the quality competition among schools, schools boasted about their students' achievement and about their high-qualified teaching staff (Greenwald et al., 1996). The more teachers possessed high certification status, for instance, a Master's degree, the more they exerted a significant impact on their students' grades, as well as their school quality. However, there are real reasons to believe that quality in education, if continuing to be measured by appraising teacher quality or school quality assurance tests, somehow, is rippled with holes (Baxter, 2014). In terms of teacher quality, it seemed that many quality of education measuring models, for example, the value-added models, only focused on the homogeneity of a sample within its context without anticipating the heterogeneity, which leads to imprecise estimation (Sass et al., 2014).

Previously, there was a lack of management know-how from educators, who mostly have been graduated in education schools (Harvey & Green, 1993). Regarding the wave of globalisation in recent decades, more and more mordent management approaches have been adopted into the educational context. For instance, Total Quality Management (TQM) expanded educational mangers' perspectives to consider the school operational process as a service and therefore, contributed to constructing the concepts of quality assurance in education (China, 2014). However, many scholars studied TQM in education and identified difficulties for schools to identify and satisfy quality elements (Hassan et al., 2013).

Obviously, teachers, students, and the school, though each is important, cannot be assessed separately when trying to reach the ultimate education objectives (Hoang et al., 2020). Moreover, although quality assessment in education, by its very nature, would contribute to the education system improvement, the notable recent movements did not demonstrate best practices for schools to refer to. For instance, international assessments, such as OECD or PISA, on the one hand, carried functional policy implications, but limited practical applications for schools (Hanushek & Woessmann, 2011), and developing countries often hold a misconception about raising their status in the PISA ranking. However, those such rankings do not contribute any aspect to the renovation at school level. At the school level, Kane and Staiger (2001) reported that more and more schools focus on test scores as a means to affirm their educational quality. This is somehow a consequence of the accountability wave. In addition, because the number is so convenient to assess, real academic progress sometimes seems to be ignored intendedly (Doherty, 2008). Specifically, their misconceptions claimed upon the international perspectives that being high grade in such competitive organisations partly asserted their insider.

Nevertheless, a better economic background had a powerful effect on the educational quality (Hanushek & Kimko, 2000). That is, evaluating the quality of an education programme by measuring the overall cognitive skills in literacy and numeracy seemed mistaken. It is difficult to determine whether students acquired these skills from formal, informal or non-formal education. Recently, governments around the world have strongly advocated developing human capital through schooling provision, an approach reflected in STEM alignment (Kataoka et al., 2020). Indeed, many developed countries, and significantly, developing countries, responded enthusiastically to integrating STEM subjects without any skepticism (Kelley & Knowles, 2016). By all accounts, it appeared patently absurd if schools defined STEM teachers' more effectively than non-STEM teachers due to its supposedly opportune orientation (Williams, 2011). In addition, when students reached remarkable achievement in this approach, it was not enough to regularise both the quality of teacher and school as well (Corlu et al., 2014).

In the following parts of this study, we first undertake an overview of prior literature on applied educational quality assessment; then propose a conceptual framework and adopt the method to collect the data; next deliver the empirical results as well as finally discuss the findings, limitations of the study and directions for future research.

LITERATURE REVIEW

Education Quality Measurement

Education quality under various prisms

Sallis (2014) defined three different concepts of "quality" over the past 800 years. The first concept, "Quality Control", came during the medieval era in Europe, when groups of tradesmen were formed (Shah et al., 2011). At that time, "quality" was primarily concerned with meeting specific standards for product, and this understanding had remained as mainstream until the early 19th century. Afterwards, the definition evolved to "Quality Assurance", due to markets blooming and industry reaching new levels of efficiency. Systems and processes tried to ensure "quality" along each step of the process by using statistics, highly detailed, and highly monitored processes.

Even though "Quality" is an ambiguous term, and it is a challenge to define (Sallis, 2014), education quality has been addressed as a significant concern across countries around the globe. Harvey and Green (1993) and Pfeffer and Coote (1991) describe quality as "a slippery concept" because numerous disparate opinions are revolving around this issue. The efforts to measure quality in education also face that same quandary. Education itself is a vast and complicated term, with so many related factors of quality in different situations (Gibbs, 2010). Determining education quality is an art, in which we will lose objectivity while evaluating (Brown, 1957). Therefore, assessing education quality is a perplexing task for researchers (Sayed & Ahmed, 2011).

Harvey and Green (1993) claim that in a democratic society, there must be enough space for each person to hold many different understandings at once. Therefore, sometimes people understand a concept in many ways, without realising that there exist some contradictions between those perspectives. For example, Green (1994) points out five distinct definitions for "quality" in higher education. However, he concludes that, of those five, only two should be considered suitable for application in higher education. Those two are: "fitness for purpose" and "effectiveness in achieving institutional goals". The openness and flexibility are reflected in relativity or subjectivity regarding education quality, which is a cause of difficulty for researchers.

As one of the first to approach quality from a "humanistic perspective", Beeby (1966) created a model of education quality based on different development stages of the characteristics of the school (curriculum and classroom) and the teacher's education and preparation. While a very Western educational viewpoint may have limited it, it is unfortunate that, at that time, more people did not take notice of the significant step forward he made by looking at the educational quality (Rayner, 2017). He looked at educational quality through a modern, progressive lens of student-centric philosophies.

In a different approach, Chitty (2002) outlined three targets of education: "human fulfilment", "preparation for the world of work" and "contributing to social progress and social change". Followed up the work of Chitty (2002) using a humanistic approach, Barrett et al. (2006) conducted an in-depth look into the critical works of many researchers, including from the World Bank, UNESCO and EFA (Education for All), and proposed five dimensions to measure quality: *Effectiveness, Efficiency, Equality, Relevance and Sustainability.* Thereafter, Nikel and Lowe (2010) added two more dimensions: *Responsiveness* and *Reflexivity* to create a new model of education quality, a piece of fabric which may be stretched in seven different directions. When one dimension is given too much emphasis, the fabric is pulled out of shape, therefore ideally each dimension is being pulled equally in all directions, achieving equilibrium, and the fabric is stretched to its limit.

As a complement to the humanistic view, Cheng and Cheung (1997) tackled education quality using management perspectives. Cheng (2003) described three waves of education quality assurance. The first wave emerged before 1980, focused on internal effectiveness. In this wave, three models emerged: Goal and Specification Model, Process Model, and Absence of Problem Model. The second wave started in the 1980s, focused on accountability to the major stakeholders. It addressed: Resource-Input Model, Satisfaction Model, Legitimacy Model, Organizational Learning Model, and Total Quality Management Model. Finally, the third and most recent wave looks at future quality and the way education is addressing the future requirements of stakeholders. This wave is notable for addressing the development of contextualised multiple intelligences and "triplisation" (Globalisation, Localisation, and Individualisation) (Cheng, 2003). Perhaps the combination of these three waves in education will bring the high quality that we are seeking, or create a paradigm shift in the understanding of education quality (Garira, 2020).

Educational service quality measurement

One of the most popular "satisfaction models" in the second wave that Cheng (2003) mentioned is called SERVQUAL (Parasuraman et al., 1988, p. 30). The model measures service quality using gap scores, which are the differences between customers' expectations and their actual perceived perception after using the service, including five dimensions: *Tangibles; Reliability; Responsiveness; Assurance; and Empathy.* Businesses can use SERVQUAL to gauge their own performance, as well as to compare with their competitors (Yousapronpaiboon, 2014). To date, SERVQUAL has been applied in many studies with different settings, from the healthcare (Muhammad Butt & Cyril de Run, 2010), retail bank (Newman, 2001), fast-food restaurants (Cronin & Taylor, 1992), to education (Cook & Thompson, 2000). As the higher education sector experienced faster internationalisation processes, it is more likely for universities and colleges to adopt SERVQUAL (Galeeva, 2016). Regarding the adoption of SERVQUAL into

K-12 education, Asubonteng et al. (1996) and Ramseook-Munhurrun and Nundlall (2013) declared the need to revise the model.

This study applies the "perceived service quality" concept of SERVQUAL, meaning the authors assess the "perceived education quality" of students via their perceptions and expectations. However, because SERVQUAL primarily deals with service quality, not all the items were utilised. There are some reasons given for why the authors chose to use students' voices as the main subjects. While there has been a great deal of research in the field that refers to the perspectives of parents (Incesu & Asikgil, 2012; Stepanova et al., 2017) and teachers (Newchurch, 2017; Hoang, 2023a) and other stakeholders (Abidin, 2015), recently there is a greater call for including students' voice and perspective into the definition and assessment of quality education (Akareem & Hossain, 2012; Garwe, 2015; Mitra, 2018). It seems self-evident that those evaluating education quality should seek students' opinions and thoughts because it is their results and outcomes that often yield much of the data used to determine the quality of education (Levin, 2000). In addition, students have at least as much, if not more, participation in the educational process as teachers have. They should be considered an important source of information. Students are even inserting themselves into the shaping of their educational experience. For example, in the state of Kentucky, a student-led organisation has inserted itself as a critical and influential voice in the state's education policy (Prichard Committee, 2013).

Approaches to identify quality measurement indicators

The economic crisis of the 1970s inspired the quality management movement, including the promotion of three conventional approaches: TQM, Performance Indicators (PIs) and External Quality Monitoring (EQM) (Chung, 2010). Around the 1950s, the Japanese broadened those Western ideas into the TQM approach, which boosts the spirit of all employees and managers to create customers' satisfaction. TQM contributed to shaping the Japanese working culture, which was crucial to the country's success after World War II. TQM has been widely applied in the field of education since the 1990s (Tsuda, 1995) and was considered a feasible solution for quality in education (Manatos et al., 2017; Sfakianaki, 2019). Researchers strived to construct a scale to measure TQM in the practice (Asif et al., 2013; Sfakianaki, 2019). Some of the indicators included things such as assessment aimed toward development; well-trained staff; a clean environment and modern infrastructure (Bayraktar et al., 2008). These indicators vary depending on the context or the goals of each institution. According to Sahney et al. (2004), TQM in education includes two essential facets: who are "customers" and the "production process". The authors referred to assorted views of determining who customers of education are. Although it received a lot of positive response, there is still some skepticism from educators concerned with its industry and business origins (Koch, 2003). One concern is there are still barriers for

which research has not found the critical solution. TQM requires a strong commitment and alignment to a unified mission by every part of the system, including the work and steps it takes to achieve it. However, achieving this type of synergy also requires a greater workload, as documentation and other procedures dictate an additional time or resource commitment on the part of all stakeholders (Abbas, 2020).

During the 1980s, educational institutions faced enormous pressure for accountability from the government and the public (Gaither, 1994). Institutions utilise PIs to prove their "value for money" (Cheng & Tam, 1997), their effectiveness concerning standards of quality. These indicators could be flexible between different schools and might include: student achievement, retention rate, class size, staff-students ratio, pass or fail rate of students. Schools believe that when they can boost their PIs, it will powerfully and positively affect teachers' behaviour (Rosenkvist, 2010); boosting education quality. However, some research shows that too much attention on PIs or public image may lead to not focusing on students' learning or even the school's actual performance (Figlio & Loeb, 2011; Hanushek & Raymond, 2005). Furthermore, some schools use "behaviour strategies" during the days that students take important exams (Anderson & Butcher, 2006; Figlio, 2006; Figlio & Winicki, 2005). Those adjustments reduce the reliability of the test results. Additionally, there are still concerns about the fairness of comparisons and whether these comparisons can accurately measure each student's development (Heck, 2000). Finally, there are arguments around the question: what are good performance indicators? (Barbato et al., 2022).

The last of the three approaches is EQM, which is carried out by outside inspectors to monitor quality through auditing, accreditation, assessment, or external examination (Harvey & Newton, 2004). EQM was put into action in various countries around the world, for instance, the US (Amaral, 1998), South African (Botha et al., 2008), Hong Kong (Wong & Li, 2010), India (Kumar et al., 2011), the Netherlands (Chu & Westerheijden, 2018). In general, EQM is based on not only the numbers but also on different qualitative criteria for assessment, such as goals, human capital, infrastructure, curriculum, or operation. Institutions using EQM have space to do self-evaluation to develop their quality assurance processes (Brennan & Shah, 2000; Smeby & Stensaker, 1999). Research shows that EQM can strengthen communication and help clarify accountability (Carroll, 1997; Dill, 2000); therefore, improving quality education. Unfortunately, it may also make accounting and compliance standards become the goal. Thus, actual quality enhancement is no longer the primary target (Harvey & Newton, 2004). Besides, EQM is not strong enough to make positive changes for some theoretical issues such as curriculum, leadership and school's climate (Horsburgh, 1999). EQM also tends to control more than improving quality. It does not always address some of the significant concerns or particular educational issues. For example, paperwork displaces teachers' time to improve their teaching; or students have to pay more, due to fees needed to cover the cost of external quality assurance (Harvey, 2006; Stensaker et al., 2011).

The Challenges of Education Quality Measurement within the Era of Globalisation

Quality in education is a subjective matter (Doherty, 2008), which can be determined as "fitness for purpose" or "fitness of purpose" (Cartwright, 2007). The quality in education, for decades, has come into a state of preoccupation for many people, especially in developing countries (Tran et al., 2019; Vuong, 2018). Unterhalter (2019) also stated that global indicators such as SDG4 follow certain political targets and did not ensure the meanings which it is representing. Ladd (1999) found the effectiveness of these programmes, which acted as an incentive for schools to be more competitive in training students for better performance gains. Wöbmann et al. (2007) also indicated from PISA 2000 that the more times students exercised standardised tests, the higher their outcomes improved. However, many argued that the education system was sticking to a deep-seated belief in the power of student achievement and fell into a false sense of the confident status quo hereafter. For instance, Kane and Staiger (2001), supposed that using tests to determine school ranking was inappropriate, and schools would suffer a loss of prestige when following a spate of rating student test scores. Also, starting the education quality by examining the effectiveness of schools within the paper and pencil form shows many limitations itself. Heck (2000) suggested that academic results reflected a combination of factors such as family background, parent academic level or teacher efficacy. Students, simultaneously, appear to be promising servility for top-down and carefully designed tests (William, 2010). Kane and Staiger (2002) expressed their disapprobation of the abuse of test-for-raking and the advanced preparation for these tests was inevitable, resulting from the imprecise measurement. Moreover, standardised tests seem likely to evoke careful arrangements, which even occurred in the nutritional assessment for school meals (Figlio & Winicki, 2005).

Furthermore, it would not be possible to assess the quality of an education program by theoretical assessments, especially when the global education system primarily emphasises the seamless integration between the very structured curriculum and the particular context (Hoang, 2017; Reimers, 2020). Razinkina et al. (2018) stated that one of the most critical shortcomings in the myriad educational quality determinations was forgetting the direct quality recipients themselves. It is obvious that student satisfaction in learning should be early attached to the quality of education assessment process. Additionally, Alghamdi (2016) indicated some obstacles such as the manager's steadfastness, fiscal problems and manpower shortage. Lee and Zuilkowski (2017) added that equality in the education system was a matter of utmost importance. According to Tikly and Barrett (2011), the inferior educational equality worked for its raison d'être: the disproportion, underprivileged, and impoverishment.

CONCEPTUAL FRAMEWORK

Considering this current void towards overall education quality in education measurement, Figure 1 presents a new quality in education measurement framework, which covers six fundamental dimensions, including quality of education criteria such as Teaching and Learning, Curriculum and Assessment, and other operational aspects like Leadership Engagement, School Climate, Community Engagement, and Sustainable Development. The proposed conceptual framework was built after reviewing relevant literature and existing frameworks, as well as consulting a group of school managers. The first dimension is the Leadership Engagement (China, 2019; Council of International Schools, n.d.; Heck, 2000; Hong Kong Education Bureau, 2016; MCIEA - Massachusetts Consortium for Innovative Education Assessments, n.d.; Van Damme, 2011). China (2019) recommended to Romanian secondary schools that school management is one of the crucial pillars of improving educational quality (besides teacher quality and educational products). In addition, many studies emphasize the role of school leaders, especially regarding problem-solving and decision-making related to school quality improvement (Heck, 2000), school objectives or student learning responsibilities (Council of International Schools, n.d.; Van Damme, 2011), educational inspection (Baxter, 2014) and management styles (Heck, 2000; MCIEA - Massachusetts Consortium for Innovative Education Assessments, n.d.). Precisely, in the school context, the central role is occupied by the middle managers, whose job is to ensure that the school's vision is realised (Terrell et al., 1996). Middle managers are also the people who have close contact with teachers, supporting and motivating teachers to tackle challenges and be proactive during a time of school change (Hong Kong Education Bureau, 2016).

Figure 1. Framework of EdQUAL - Overall Education Quality Measurement scale

The second dimension is Teaching and Learning (Hong Kong Education Bureau, 2016; Hoy et al., 2006; MCIEA - Massachusetts Consortium for Innovative Education Assessments, n.d.). Introduced in 2003, the School Development and Accountability (SDA) framework has often been used for school self-evaluation in Hong Kong. There are 23 performance indicators for implementing the framework and six of that focus on teaching and learning domains. Teachers should teach in an understandable and well-organised manner, with clear instructions and demonstration (MCIEA - Massachusetts Consortium for Innovative Education Assessments, n.d.), while "creating a lively classroom learning atmosphere with good class discipline" (Hong Kong Education Bureau, 2016). Teachers should also encourage students to set high learning achievement goals (MCIEA - Massachusetts Consortium for Innovative Education Assessments, n.d.), since it will enhance optimism and the collective efficacy of the classroom, thus motivating teachers and students "to act to accomplish challenging goals and persist until they are successful" (Hoy et al., 2006). Moreover, students should receive support promptly, both inside and outside school, on both academic and personal problems, so that they can apply what they have learnt, reflect on their performance and gain experience as well as self-management skills (Hong Kong Education Bureau, 2016).

The third dimension is the Curriculum and Assessment (Council of International Schools, n.d.; Hong Kong Education Bureau, 2016; Rust, 2002; Stabback, 2016). According to the UNESCO-IBE framework conceptualised by Stabback (2016), a school's curriculum should be inclusive and integrated across all subjects. Council of International Schools (n.d.) also agreed to this criterion, claiming that a curriculum should consist of real-life content such as digital citizenship, global citizenship, and intercultural learning to help develop students' knowledge, understanding, skills, attributes, well-being and future preparation. In addition, schools should have a timeline to revise and evaluate the curriculum on a regular basis (Hong Kong Education Bureau, 2016). Furthermore, curriculum content can also be assessed through students' learning process, but schools need to ensure assessment criteria are concrete and clear (Rust, 2002).

The fourth dimension is School Climate (Council of International Schools, n.d.; Heck, 2000; Pulis, 2018; Zedan, 2010). Heck (2000) contended that an important school indicator is a positive school climate, which can be demonstrated by "a safe environment, clean and comfortable buildings, and teachers demonstrate caring attitudes". Pulis (2018) concurred, finding that, according to student surveys, a "good" school has a safe environment, pleasant atmosphere, and trust and respect between pupils and teachers (Zedan, 2010). Additionally, it has been suggested that school environments should represent mutual respect, openness, fairness and trust to reinforce student learning related to well-being while increasing their desire to study and motivating them to learn more (Council of International Schools, n.d.). Also, the changes in macro factors like the recent impacts from COVID-19 also affect quality of education across national and organisational levels (Ramrathan, 2021).

The fifth dimension is Community Engagement (Beabout and Jakiel, 2011; Hong Kong Education Bureau, 2016; Odendaal & Plessis, 2018; Tikly & Barrett, 2011). In Tikly and Barrett's study (2011), a good quality education would involve the home and community environments. In particular, parent-teacher relationships are associated with students' academic engagement and achievement (Hughes & Kwok, 2007); and schools in Hong Kong consider parents as essential partners in school development (Hong Kong Education Bureau, 2016). Furthermore, external educational institutions and experts can also be helpful for schools' quality improvement. Specifically, external consultants can be the reasons for schools' innovation, improvement and effectiveness (Odendaal & Plessis, 2018); some schools even used this strategy to turn around low-performing situations (Beabout & Jakiel, 2011). Regarding external institutions, many studies confirmed community engagement could provide resources and opportunities for students, such as promoting students' social, emotional and academic learning (Mahoney et al., 2021).

The final dimension is Sustainable Development (Hodson & Sander, 2017; Lagrosen et al., 2004; MCIEA, n.d.; Nikel & Lowe, 2010; Sakthivel et al., 2005). Among seven dimensions of school quality, Nikel and Lowe (2010) defined the sustainability dimension as "taking up responsibility for global environmental changes and the uncertainty over the well-being of future generations". This sustainability dimension can have a considerable impact on school quality. In particular, campus facilities can predict the students' satisfaction (Lagrosen et al., 2004; Sakthivel et al., 2005), whilst the number of trees covered in student environments can enhance academic success (Hodson & Sander, 2017). What is more, investing in infrastructure, equipment and learning resources can significantly affect the achievement of students (Lagrosen et al., 2004; MCIEA, n.d.), especially disadvantaged learners (Tikly & Barrett, 2011). Last but not least, professional development is of great importance in improving school effectiveness. Schools should promote professional exchange and teacher reflection on their own work so that teachers' instructional practices, skills and experience are kept up-to-date (Heck, 2000; Hong Kong Education Bureau, 2016) and so the school's learning community is established. As a result, teachers' professional development will contribute to "the development of student learning, well-being, and global citizenship" (Council of International Schools, n.d.). In addition to the mentioned literature, the authors also studied documents from the Vietnam Ministry of Education and Training (MoET) including the "Regulation on education quality accreditation and recognition of national standards for secondary schools and inter-level schools" (Vietnam Ministry of Education and Training, 2018b). The document mentions all six proposed dimensions with items related to teachers' professional development, management competencies, safe and green environment, learning facilities, school-parent relationships, effective assessments, and integrated curriculum.

Regarding the scope of this research, the further sections will confirm the validity and reliability of the scale with Cronbach's alpha and Confirmatory Factor Analysis (CFA).

Finally, we will apply this scale to measure quality in eight lower secondary schools in Ho Chi Minh City. Following the well-known methodology of SERVQUAL (Parasuraman et al., 1988), the research group will use the scales to measure the mean gap differences between students' expectations and perceptions about overall education quality.

METHODOLOGY

This research uses quantitative analyses to validate the proposed conceptual framework for overall education quality measurement. First, we implemented a pilot study including 120 randomly selected students from two schools and eliminated all irrelevant, ambiguous or similar items. As a result, the final version of the EdQUAL scale (see Table 1) consists of 19 items and six dimensions (3 to 4 items in each dimension). For each question, students respond to a 5-point Likert scale with the endpoints 1 = Totally disagree to 5 = Totally agree.

Table 1. EdQUAL's dimensions of overall quality in education

Dimensions	Items	
Leadership engagement	1.	The school leaders help students to understand the school's goals.
	2.	The school leaders deal well with all problems in school.
	3.	The school leaders have stringent management in order to enhance the quality of each lesson.
Teaching and learning	1.	Teachers' instruction is clear and understandable.
	2.	Students receive support in a timely manner.
	3.	Students are encouraged to set high learning achievement goals.
Curriculum and assessment	1.	The curriculum is well integrated.
	2.	The curriculum is revised on a regular basis.
	3.	The assessment criteria are concrete and clear.
School climate	1.	The school environment is comfortable.
	2.	The school makes everyone feel safe.
	3.	The school communicates well with external educational experts.

(Continued on next page)

Table 1: (*Continued*)

Dimensions	Items
Community engagement	1. The school communicates well with parents. 2. The school communicates well with other educational institutions. 3. The school communicates well with external educational experts.
Sustainable development	1. Teachers are often involved in continuous professional development activities. 2. The school has modern, adequate learning resources and equipment. 3. The school's campus is spacious, airy and full of greenery.

Vietnam, for many centuries, had been influenced by Confucianism, which means government officers were selected based on academic results, and many education-related processes were centralised by the Vietnamese State government. Nonetheless, Vietnam witnessed a trend of decentralising the education system, across the system of 28,922 K-12 schools (Hoang et al., 2020). Specifically, schools have more autonomy to choose textbooks based on the common standard curriculum (Vietnam Ministry of Education and Training, 2018a). However, the schools' quality measurement or ranking is still implemented by the Department of Education and Training and the procedure does not really take into account students' opinions (Vietnam Ministry of Education and Training, 2006). Our framework starts with the viewpoints of students and can be the foundation to expand to other stakeholders.

Data were collected from eight lower-secondary schools in Ho Chi Minh City from March to June 2020 under the clustering sampling method, including four public schools, two private schools, and two international schools. The schools were established between 1964 and 2020. Most of the schools follow the curriculum of Vietnam MoET, the rest follow the Cambridge program or combine the Vietnam curriculum with a foreign curriculum. These eight schools cover three types of schools, have distinctive management models and implement different curricula; hence we can expect generalised results for other Vietnamese schools outside the sample.

The research protocol was approved by our institutional IRB in February 2020 (No. 260220). Upon the agreement of principals and teachers from each participating school, researchers delivered the consent forms indicating participants' confidentiality, anonymity and voluntary status to students. The teachers then gave survey links to students. The first round of the survey to record students' expectations when they started learning at their current schools was sent first, and then the second round to

record students' perceptions was sent two weeks later. This process was replicated in all schools, and finally, 2,292 observations were collected in total. After cleaning the dataset, 2,239 observations were valid for analysis, and the full version of the dataset can be found on the Harvard Dataverse repository (Hoang, 2023b).

RESULT

Descriptive Statistics

Table 2 illustrates the statistic of the participants' demographic. The surveyed students are distributed equally regarding grade and gender. Most of them (94.5%) started learning in the surveyed schools from grade 6. Thus, they tend to understand their schools adequately. In addition, 87.5% of students study in class with 36 to 45 students.

Table 2. Demographic data of the sample (N = 2,239)

Sample	Frequency	%	Valid %	Cumulative %
Grade				
6	567	25.3	25.3	25.3
7	426	19.0	19.0	44.4
8	659	29.4	29.4	73.8
9	587	26.2	26.2	100.0
Gender				
Female	1,174	52.4	52.4	52.4
Male	1,065	47.6	47.6	100.0
Start learning at this school since…				
Grade 6	2,116	94.5	94.5	94.5
The beginning of this academic year	123	5.5	5.5	100.0
Class size				
13–24 students	66	2.9	2.9	2.9
25–35 students	214	9.6	9.6	12.5
36–45 students	1,959	87.5	87.5	100.0

Reliability and Validity Analysis

The dataset was analysed with Statistical Package for the Social Sciences (SPSS) Software version 20. First, we want to check the validity of the EdQUAL. Drost (2011) claimed that if research components are meaningful, then the research validity is high.

Researchers in this study want to know whether EdQUAL scale measures quality in education. However, these kinds of questions can never be perfectly answered; thus researchers need to develop persuasive reasons to support their study's validity (Bollen, 2014).

Among four different types of validity (Drost, 2011), we consider internal validity in this study as it is considered the "*sine qua non* of meaningful research" (Pedhazur & Schmelkin, 2013, p. 224) and "internal validity speaks to the validity of the research itself" (Drost, 2011, p. 115). According to (Miles & Huberman, 1994), internal validity brings into three questions: "Do the findings of the study make sense? Are they credible to the people we study and to our readers? Do we have an authentic portrait of what we were looking at?" (p. 278).

Besides employing a panel of experts for developing theoretical construct and implementing a pilot study, we use CFA to confirm the validity of the EdQUAL. First, we run the Kaiser-Meyer-Olkin (KMO) Test to assure the suitability for factor analysis. The KMO value is 0.833 for expectation's items and 0.764 for perception's items, indicating meritorious and middling results, respectively (Kaiser, 1974) (see Appendix A).

Next, we use CFA to examine if EdQUAL associates all collinear variables to the same latent variable regarding their common correlation (Morinaj et al., 2017). From the loadings in the pattern matrix tables (Appendix B and C), it is clear that all items are divided into six factors as proposed in our framework. Next, the most important indices of model fit (Schermelleh-Engel et al., 2003) are reported in Table 3, including the chi-square test statistic (χ^2), the relative/normed chi-square (χ^2/df), the root mean square error of approximation (RMSEA), the standardised root mean square residual (SRMR), the adjusted goodness-of-fit statistic (AGFI), the comparative fit index (CFI) and the Parsimonious normed fit index (PNFI). All the model fit indices reach adequacy level regarding the following criteria: χ^2/df < 5 (Wheaton et al., 1977), RMSEA < 0.08 (MacCallum et al., 1996), SRMR < 0.05 (Byrne, 2013), AGFI > 0.90, CFI > 0.95 (Hu & Bentler, 1999) and PNFI > 0.5 (Mulaik et al., 1989). As a result, the proposed EdQUAL framework is a well-fitting one.

Table 3. Goodness of fit indices for CFA model

Model	No. of item	(χ^2)	df	(χ^2/df)	RMSEA	SRMR	AGFI	CFI	PNFI
Expectation**	19	554.583*	118	4.700	0.041	0.0361	0.959	0.983	0.675
Perception**	19	673.531*	137	4.916	0.042	0.0381	0.958	0.957	0.759

Note: *Significant at p < 0.001 and n > 500, therefore χ^2/df value is referred (Wheaton et al., 1977). **Factor indicates to latent constructs in this research.

In addition, researchers examine the reliability of the EdQUAL scale by computing Cronbach's alpha (Cronbach, 1951), the most popular method for testing internal consistency (Drost, 2011). We calculated the Cronbach's alpha for each dimension, and the results range from 0.638 to 0.908, which are considered from "acceptable" to "strong" (Taber, 2018). The full results of reliability analysis are shown in Table 4 and Table 5, regarding students' expectations and perceptions, respectively. All in all, the EdQUAL is proved to be reliable and validated.

Table 4. Reliability analysis result regarding students' expectations

Expectation of quality in education					
Dimension	Scale			Reliable statistics	
	Scale mean	Item means	Item variances	Cronbach's alpha	*N*
Leadership Engagement	12.13	4.042	0.625	.892	3
Teaching and Learning	12.06	4.021	0.648	.908	3
Curriculum and Assessment	12.26	4.085	0.608	.791	3
School Climate	16.05	4.013	0.568	.840	4
Community Engagement	12.18	4.060	0.564	.799	3
Sustainable Development	12.07	4.022	0.533	.795	3

Table 5. Reliability analysis result regarding students' perceptions

Perception of quality in education					
	Scale			Reliable statistics	
Dimension	Scale mean	Item means	Item variances	Cronbach's alpha	*N*
Leadership Engagement	11.10	3.699	0.985	.638	3
Teaching and Learning	10.60	3.534	0.798	.860	3
Curriculum and Assessment	10.46	3.488	0.906	.706	3
School Climate	13.10	3.274	0.657	.798	4
Community Engagement	9.59	3.198	0.657	.742	3
Sustainable Development	11.10	3.698	0.588	.755	3

Overall Quality in Education Gaps

The second primary part of our study is to illustrate a practical case using data collected from eight secondary schools in Vietnam. First, we use a paired *t*-test to examine the difference between students' expectations and perceptions about the schools. The results are shown in Appendices D, E, and F. From data in Appendix D, we can see

that students expected most (mean = 4.19) that they would be cared, respected, and expected least (mean = 3.79) that the school would make them love studying and want to learn more. At the moment, their perceptions give the highest mark for the school's campus (mean = 3.87) and the lowest mark for school communication with experts (mean = 3.02).

Appendix E shows the correlation statistic between each couple of items (expectation versus perception item). The results show a significant correlation among all couples of items ($p < 0.001$), and the numbers range from low to moderate (from 0.205 to 0.676). Appendix F reports the mean gap differences between perception and expectation of students about quality in education. All scores are statistically significant ($p < 0.001$) and are negative, indicating that expectation had not been met, and the quality is considered low or unsatisfied (Yousapronpaiboon, 2014). Also, the larger (while negative) the mean gap difference, the lower the quality that is measured (Lewis, 1993). It is clear from Appendix F that in general, students expect a higher quality in education from school than what they perceive recently. The smallest mean gap difference (mean gap difference = –0.261) suggests that the quality point of the school board's management is excellent, as the students' expectation and current perception is different only slightly. In contrast, the communication between schools and experts has the highest mean gap difference (–1.034), indicating an inferior quality score of communication.

Regarding the dimensions of the scale, the Sustainable Development has the highest quality score as the mean gap differences for all three items are small (namely –0.324, –0.301, and –0.346). Conversely, the Community Engagement has the lowest quality score, which means the schools do not communicate well with parents, experts and other educational institutions, with all mean gap differences are big (namely –0.784, –0.767, and –1.034).

In addition, we fed the data into Excel and plotted each school's mean values for six dimensions. The results for two noticeable schools are presented in visualised radar charts (Figure 2). It is clear from the figures that two schools witness a decrease when it comes from students' expectations to their perception of the school's quality in education. However, in general, School H has a higher quality score than School A according to students' evaluation.

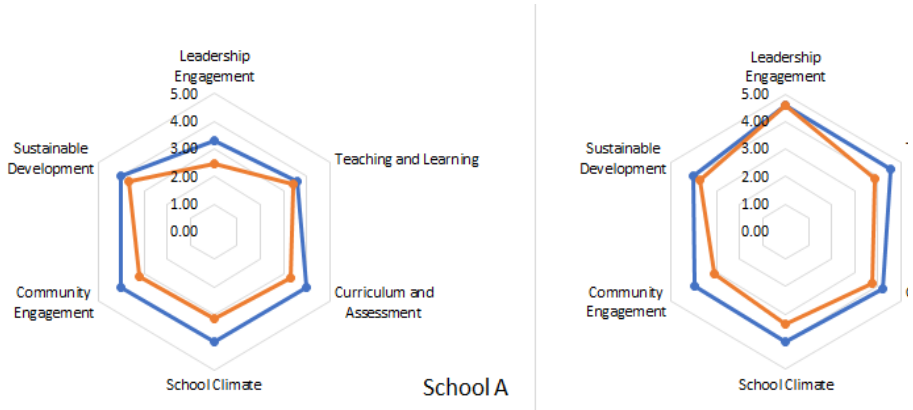

Figure 2. Sample comparison of overall quality in education

DISCUSSION

This study has introduced a lean and concrete framework to measure the overall 'quality in education', as a supplement to the quality curriculum framework under SDG 4 (Hoang et al., 2020; Stabback, 2016), which focused on 'quality of education'. Moreover, with EdQUAL, schools can not only gauge their quality in education using students' current perceptions but also detect the quality gap by assessing students' expectations and perceptions. In our research, we collected data from both students' expectations and perceptions to increase the robustness of the scale. The results from Cronbach's alpha and CFA model confirm that the scale achieves reliability and validity.

Besides presenting EdQUAL as a measurement scale, our study also harmonised the approach of SERVQUAL (Parasuraman et al., 1988) and TQM and applied in K-12 education context. Regarding the empirical of Vietnamese lower secondary schools, the mean gap differences (equals perception minus expectation) can reflect the good, the bad and the shortage among various quality in education dimensions. The most positive mean gap difference would indicate the highest education quality, and the most negative mean gap difference would imply the worst education quality (Yousapronpaiboon, 2014). All studied schools show negative mean gap differences in all six dimensions, indicating that on average, students' expectations exceed their perceptions. Thus, schools need to improve their quintessential dimensions to enhance overall quality in education. This phenomenon is also similar to the empirical results within the education sector (Tan & Kek, 2004; Yousapronpaiboon, 2014) as well as in other sectors such as health (Muhammad Butt & deRun, 2010), restaurant service (Heung et al., 2000), or banking sector (Newman, 2001). Specifically, Community Engagement has the most negative mean gap difference may be due to the limitation

in external activities for students (with experts and other institutions), or with the fact that parents do not know much about their children's learning. Therefore, schools should focus on improving communication means and tools to facilitate relationships with stakeholders, as suggested by Swick and Bailey (2004) and Pauley and Pauley (2009). In contrast, the fact that Sustainable Development has the least negative mean gap difference is highly likely due to the recent regulations in Vietnam which promoting SDG 4 such as Action Plan on Gender Equality in Education 2016–2020 (Vietnam Ministry of Education and Training, 2006), Decision No 2161 about the *Implementation of Sustainable Development Goals in Education and Training until 2025 and Orientations Up to 2030* (Vietnam Ministry of Education and Training, 2017), and especially *Vietnam's New General Educational Curriculum* (Vietnam Ministry of Education and Training, 2018a). Still, schools need to find a way to turn this negative number to positive ones.

CONCLUSION, APPLICATION, LIMITATIONS, AND FUTURE RESEARCH

The main aim of this study is to provide a distinctive framework to measure overall quality in education, which contributes to extend the impacts of prior works (Hoang et al., 2020; Stabback, 2016) towards sustainable education. History witnessed a remarkable amount of quality measurement in various sectors; still, there is no specific framework for K-12 education quality. As a result, EdQUAL is developed to provide a holistic portrait in assessing overall quality in education, in which various educational dimensions are taken into account, namely Leadership and Engagement, Teaching and Learning, Curriculum and Assessment, School Climate, Community Engagement, and Sustainable Engagement. The second aim of this research is to apply EdQUAL into secondary schools' context in Vietnam. The scales are validated, reliable, and show informative results when measuring various Vietnamese schools. Regarding EdQUAL's rigorous and robustness, the authors are confident that, it is more convenience for school leaders to capture a picture of their school's 'quality in education' with less time and resources. It is also easier to understand current gaps with EdQUAL, thenceforth proposing and monitoring renovations would be handy. Finally, individual school and school districts can adopt EdQUAL to measure their current status, whilst regional and national education leaders can consider it as a contributing index for their master renovation agenda. The measurement results using EdQUAL on a regular basis will also generate a time-series dataset, which can provide meaningful insights about education reforms.

Nevertheless, our research has some limitations. Firstly, the empirical data was collected from lower-secondary schools in Vietnam only, which might narrow the contextual nature of the study. In general, the management model of lower-secondary schools in Vietnam is not considerably different from that of upper-secondary schools.

However, students in upper-secondary schools normally have higher pressure due to the university entrance exam, and this can affect various items regarding leadership, teaching, learning or school climate within EdQUAL measurements. Thus, future research can adopt EdQUAL for various school levels (primary, upper secondary, and post-secondary) and in different schools in the world to increase scale reliability. Secondly, the collected data was completely from schools in Ho Chi Minh city, which is representative for only urban areas in Vietnam but not for rural areas. Those schools consist of public schools with not-expensive tuition fee, private schools for middle class people, and international schools for people in high class. Therefore, future studies can use the scale to measure schools in rural areas as those schools may provide unexpected results. Thirdly, the measurement scale construction process included a literature review and focus groups of school managers only. The proposed framework can also be strengthened by adding perspectives of teachers, parents, and policymakers. Fourthly, as our theoretical model focused on measuring 'quality in education', it is not yet able to measure other educational outcome-related indicators. Therefore, we expect that future studies will enable new pathways to describe the conceptual definitions of "educational products" and its measurement indicators. Last but not least, surveyed people were only students, while teachers and parents' viewpoints should be considered. Additional perspectives from teachers and parents would lead to various mean gap differences; yet provide more comprehensive and diverse insights to educators. Regarding those future improvements, we hope that researchers will publicise the datasets of their future studies to contribute to the sustainable application of this framework, as suggested by (Vuong, 2020).

REFERENCES

Abbas, J. (2020). HEISQUAL: A modern approach to measure service quality in higher education institutions. *Studies in Educational Evaluation, 67*, 100933. https://doi.org/10.1016/j.stueduc.2020.100933

Abidin, M. (2015). Higher education quality: Perception differences among internal and external stakeholders. *International Education Studies, 8*(12), 185. https://doi.org/10.5539/ies.v8n12p185

Akareem, H. S., & Hossain, S. S. (2012). Perception of education quality in private universities of Bangladesh: A study from students' perspective. *Journal of Marketing for Higher Education, 22*(1), 11–33. https://doi.org/10.1080/08841241.2012.705792

Alghamdi, H. (2016). Toward better understanding of total quality management (TQM). *Journal of Business & Economic Policy, 3*(4), 29–37.

Amaral, A. M. S. C. (1998). The US accreditation system and the CRE's quality audits: A comparative study. *Quality Assurance in Education, 6*(4), 184–196. https://doi.org/10.1108/09684889810242173

Anderson, P. M., & Butcher, K. F. (2006). Reading, writing, and refreshments. *Journal of Human Resources, XLI*(3), 467–494. https://doi.org/10.3368/jhr.XLI.3.467

Armstrong, P., Hughes, B., Courtney, S., Gunther, H., & Gardner-Mctaggart, A. (2021). *Bubble bubble toil and trouble: School leadership as Shakespearean tragedy in and after COVID-19*. British Education Research Association (BERA).

Asif, M., Awan, M. U., Khan, M. K., & Ahmad, N. (2013). A model for total quality management in higher education. *Quality & Quantity, 47*(4), 1883–1904. https://doi.org/10.1007/s11135-011-9632-9

Asubonteng, P., McCleary, K. J., & Swan, J. E. (1996). SERVQUAL revisited: A critical review of service quality. *Journal of Services Marketing, 10*(6), 62–81. https://doi.org/10.1108/08876049610148602

Barbato, G., Bugaj, J., Campbell, D. F. J., Cerbino, R., Ciesielski, P., Feliks-Długosz, A., Milani, M., & Pausits, A. (2022). Performance indicators in higher education quality management of learning and teaching: Lessons from a benchlearning exercise of six European universities. *Quality in Higher Education, 28*(1), 82–105. https://doi.org/10.1080/13538322.2021.1951456

Barrett, A. M., Chawla-Duggan, R., Lowe, J., Nikel, J., & Ukpo, E. (2006). The concept of quality in education: A review of the 'international' literature on the concept of quality in education. EdQual Working Paper 3.

Baxter, J. A. (2014). An independent inspectorate? Addressing the paradoxes of educational inspection in 2013. *School Leadership & Management, 34*(1), 21–38. https://doi.org/10.1080/13632434.2013.856294

Bayraktar, E., Tatoglu, E., & Zaim, S. (2008). An instrument for measuring the critical factors of TQM in Turkish higher education. *Total Quality Management & Business Excellence, 19*(6), 551–574. https://doi.org/10.1080/14783360802023921

Beabout, B. R., & Jakiel, L. B. (2011). Family and community engagement in charter schools. In S. Redding, M. Murphy, & P. Sheley (Eds.). *Handbook on family and community engagement* (pp. 147–152). Academic Development Institute.

Beeby, C. E. (1966). *The quality of education in developing countries.* Harvard University Press. https://doi.org/10.4159/harvard.9780674188198

Bollen, K. A. (2014). Measurement models: The relation between latent and observed variables. In *Structural equations with latent variables* (pp. 179–225). Wiley. https://doi.org/10.1002/9781118619179.ch6

Botha, J., Favish, J., & Stephenson, S. (2008). Institutional audits: A comparison of the experiences of three South African universities. *Quality in Higher Education, 14*(1), 29–53. https://doi.org/10.1080/13538320802054524

Brennan, J., & Shah, T. (2000). *Managing quality in higher education: An international perspective on institutional assessment and change*. Buckingham, UK: OECD,SRHE and Open University Press.

Brown, S. (1957). Quality in education. *Journal of Educational Sociology, 30*(8), 361. https://doi.org/10.2307/2264767

Byrne, B. M. (2013). *Structural equation modeling with Lisrel, Prelis, and Simplis*. Psychology Press. https://doi.org/10.4324/9780203774762

Carroll, M. L. (1997). Self-audits of quality in academic and service departments: A practical experience. *Assessment & Evaluation in Higher Education, 22*(2), 225–231. https://doi.org/10.1080/0260293970220210

Cartwright, M. J. (2007). The rhetoric and reality of "quality" in higher education. *Quality Assurance in Education, 15*(3), 287–301. https://doi.org/10.1108/09684880710773174

Cheng, Y. C., & Cheung, W. M. (1997). Multi-models of education quality and multi-levels of self-management in schools. *Educational Management & Administration, 25*(4), 451–462. https://doi.org/10.1177/0263211X97254008

Cheng, Y. C., & Tam, M. -W. (1997). Multi-models of quality in education. *Quality Assurance in Education, 5*(1), 22–31. https://doi.org/10.1108/09684889710156558

Cheng, Y. C. (2003). Quality assurance in education: internal, interface, and future. *Quality Assurance in Education, 11*(4), 202–213. https://doi.org/10.1108/09684880310501386

China, R. (2019). Quality assurance in secondary education: 'A Bridge Too Far'? In E. Soare, & C. Langa (Eds.), *Education Facing Contemporary World Issues, vol 67. European Proceedings of Social and Behavioural Sciences* (pp. 318–327). Future Academy. https://doi.org/10.15405/epsbs.2019.08.03.38

China, R. (2014). Calitate în educaþie versus calitatea educaþiei (I). CALITATEA, 15(139), 18–28.

Chitty, C. (2002). *Understanding schools and schooling.* Psychology Press. https://doi.org/10.4324/9780203205006

Chu, A., & Westerheijden, D. F. (2018). Between quality and control: what can we learn from higher education quality assurance policy in the Netherlands. *Quality in Higher Education, 24*(3), 260–270. https://doi.org/10.1080/13538322.2018.1559513

Chung, S. L. D. (2010). Quality assurance in post-secondary education. *Quality Assurance in Education, 18*(1), 64–77. https://doi.org/10.1108/09684881011016007

Cook, C., & Thompson, B. (2000). Reliability and validity of servqual scores used to evaluate perceptions of library service quality. *The Journal of Academic Librarianship, 26*(4), 248–258. https://doi.org/10.1016/S0099-1333(00)00114-2

Corlu, M. S., Capraro, R. M., & Capraro, M. M. (2014). Introducing STEM education: Implications for educating our teachers for the age of innovation. *Education and Science, 39*(171), 74–85. https://doi.org/10.15390/ES.2014.1219

Council of International Schools (n.d.). Membership standards. https://www.cois.org/for-schools/membership-standards

Craig, I. (2021). Whatever happened to educational management? The case for reinstatement. *Management in Education, 35*(1), 52–57. https://doi.org/10.1177/0892020620962813

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https://doi.org/10.1007/BF02310555

Cronin, J. J., & Taylor, S. A. (1992). Measuring service quality: A reexamination and extension. *Journal of Marketing, 56*(3), 55–68. https://doi.org/10.1177/002224299205600304

Dill, D. D. (2000). Capacity building as an instrument of institutional reform: Improving the quality of higher education through academic audits in the UK, New Zealand, Sweden, and Hong Kong. *Journal of Comparative Policy Analysis: Research and Practice, 2*(2), 211–234. https://doi.org/10.1080/13876980008412643

Doherty, G. D. (2008). On quality in education. *Quality Assurance in Education, 16*(3), 255–265. https://doi.org/10.1108/09684880810886268

Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and Perspectives, 38*(1), 105–123. https://doi.org/10.70953/ERPv38.11005

Figlio, D., & Loeb, S. (2011). School accountability. In *Handbook of the economics of education* (pp. 383–421). Elsevier. https://doi.org/10.1016/B978-0-444-53429-3.00008-9

Figlio, D. N. (2006). Testing, crime and punishment. *Journal of Public Economics, 90*(4–5), 837–851. https://doi.org/10.1016/j.jpubeco.2005.01.003

Figlio, D. N., & Winicki, J. (2005). Food for thought: the effects of school accountability plans on school nutrition. *Journal of Public Economics, 89*(2–3), 381–394. https://doi.org/10.1016/j.jpubeco.2003.10.007

Gaither, G. (1994). *Measuring up: The promises and pitfalls of performance indicators in higher education*. Washington, DC: Wiley.

Galeeva, R. B. (2016). SERVQUAL application and adaptation for educational service quality assessments in Russian higher education. *Quality Assurance in Education, 24*(3), 329–348. https://doi.org/10.1108/QAE-06-2015-0024

Garira, E. (2020). A proposed unified conceptual framework for quality of education in schools. *SAGE Open, 10*(1), 215824401989944. https://doi.org/10.1177/2158244019899445

Garwe, E. C. (2015). Student voice and quality enhancement in higher education. *Journal of Applied Research in Higher Education, 7*(2), 385–399. https://doi.org/10.1108/JARHE-05-2014-0055

Gibbs, G. (2010). *Dimensions of quality.* Higher Education Academy (UK).

Green, D. (1994). *What is quality in higher education?* Taylor & Francis.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66*(3), 361. https://doi.org/10.2307/1170528

Guyton, E., & Farokhi, E. (1987). Relationships among academic performance, basic skills, subject matter knowledge, and teaching skills of teacher education graduates. *Journal of Teacher Education, 38*(5), 37–42. https://doi.org/10.1177/002248718703800508

Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American Economic Review, 90*(5), 1184–1208. https://doi.org/10.1257/aer.90.5.1184

Hanushek, E. A., & Raymond M. E. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management, 24*(2), 297–327. https://doi.org/10.1002/pam.20091

Hanushek, E. A., & Woessmann, L (2011). How much do educational outcomes matter in OECD countries? *Economic Policy, 26*(67), 427–491. https://doi.org/10.1111/j.1468-0327.2011.00265.x

Harvey, L. (2006). Impact of quality assurance: Overview of a discussion between representatives of external quality assurance agencies. *Quality in Higher Education, 12*(3), 287–290. https://doi.org/10.1080/13538320601051010

Harvey, L., & Green, D. (1993). Defining quality. Assessment & Evaluation in Higher Education, 18(1), 9–34. https://doi.org/10.1080/0260293930180102

Harvey, L., & Newton, J. (2004). Transforming quality evaluation. *Quality in Higher Education, 10*(2), 149–165. https://doi.org/10.1080/1353832042000230635

Hassan, A., Ip-Shing, F., & Johnstone, A. (2013). Comparison between TQM CSFS in service sector and education sector [Paper presentation]. ICERI2013 Proceedings, Seville, Spain, 5906–5915. IATED.

Heck, R. H. (2000). Examining the impact of school quality on school outcomes and improvement: A value-added approach. *Educational Administration Quarterly, 36*(4), 513–552. https://doi.org/10.1177/00131610021969092

Heung, V. C. S., Wong, M. Y., & Qu, H. (2000). Airport-restaurant service quality in Hong Kong. *Cornell Hotel and Restaurant Administration Quarterly, 41*(3), 86–96. https://doi.org/10.1177/001088040004100320

Hoang, A. -D. (2023a). A bibliometrics analysis of research on teachers' satisfaction from 1956 to 2022. *International Journal of Educational Management, 37*(1), 164–185. https://doi.org/10.1108/IJEM-01-2022-0009

Hoang, A. -D. (2023b). Replication data for Vietnamese students' expectation and perception on educational quality. *Harvard Dataverse*. https://doi.org/10.7910/DVN/9S1AX2

Hoang, A. -D., Pham, H. -H., Nguyen, Y. -C., Nguyen, L. -K. -N., Vuong, Q. -H., Dam, M. Q., Tran, T., & Nguyen, T. -T. (2020). Introducing a tool to gauge curriculum quality under Sustainable Development Goal 4: The case of primary schools in Vietnam. *International Review of Education, 66*(4), 457–485. https://doi.org/10.1007/s11159-020-09850-1

Hoang, L. (2017). Accountability in Vietnam's education: Toward effective mechanism in the decentralization context. Country case study prepared for the 2017/8 Global Education Monitoring Report, UNESCO.

Hodson, C. B., & Sander, H. A. (2017). Green urban landscapes and school-level academic performance. *Landscape and Urban Planning, 160*, 16–27. https://doi.org/10.1016/j.landurbplan.2016.11.011

Hong Kong Education Bureau. (2016). Performance indicators for Hong Kong schools for secondary, primary and special schools. https://www.edb.gov.hk/en/sch-admin/sch-quality-assurance/performance-indicators/index.html

Horsburgh, M. (1999). Quality monitoring in higher education: the impact on student learning. *Quality in Higher Education, 5*(1), 9–25. https://doi.org/10.1080/1353832990050102

Hoy, W. K., Tarter, C. J., & Hoy, A. W. (2006). Academic optimism of schools: A force for student achievement. *American Educational Research Journal, 43*(3), 425–446. https://doi.org/10.3102/00028312043003425

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hughes, J., & Kwok, O. (2007). Influence of student-teacher and parent-teacher relationships on lower achieving readers' engagement and achievement in the primary grades. *Journal of Educational Psychology, 99*(1), 39–51. https://doi.org/10.1037/0022-0663.99.1.39

Incesu, G., & Asikgil, B. (2012). An evaluation of the relationship between service quality in primary education and parent satisfaction. *International Journal of Business and Management, 7*(18), 92–98. https://doi.org/10.5539/ijbm.v7n18p92

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31–36. https://doi.org/10.1007/BF02291575

Kane, T., & Staiger, D. (2001). Improving school accountability measures. NBER Working Paper No. 8156. https://doi.org/10.3386/w8156

Kane, T. J., & Staiger, D. O. (2002). The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives, 16*(4), 91–114. https://doi.org/10.1257/089533002320950993

Kataoka, S., Vinh, L. A., Kitchlu, S., & Keiko, I. (2020). *Vietnam's human capital: Education success and future challenges*. World Bank. https://doi.org/10.1596/34316

Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education, 3*(1), 11. https://doi.org/10.1186/s40594-016-0046-z

Koch, J. V. (2003). TQM: Why is its impact in higher education so small? *The TQM Magazine, 15*(5), 325–333. https://doi.org/10.1108/09544780310487721

Kumar, R., Garg, D., & Garg, T. K. (2011). TQM success factors in North Indian manufacturing and service industries. *The TQM Journal, 23*(1), 36–46. https://doi.org/10.1108/17542731111097470

Ladd, H. F. (1999). The Dallas school accountability and incentive program: An evaluation of its impacts on student outcomes. *Economics of Education Review, 18*(1), 1–16. https://doi.org/10.1016/S0272-7757(97)00044-7

Lagrosen, S., Seyyed-Hashemi, R., & Leitner, M. (2004). Examination of the dimensions of quality in higher education. *Quality Assurance in Education, 12*(2), 61–69. https://doi.org/10.1108/09684880410536431

Lee, J., & Zuilkowski, S. S. (2017). Conceptualising education quality in Zambia: A comparative analysis across the local, national and global discourses. *Comparative Education, 53*(4), 558–577. https://doi.org/10.1080/03050068.2017.1348020

Leu, E., & Price-Rom, A. (2006). *Quality of education and teacher learning: A review of the literature*. U.S. Agency for International Development.

Levin, B. (2000). Putting students at the centre of educational reform. *Journal of Educational Change, 1*, 155–172. https://doi.org/10.1023/A:1010024225888

Lewis, B. R. (1993). Service quality measurement. *Marketing Intelligence & Planning, 11*(4), 4–12. https://doi.org/10.1108/02634509310044199

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130

Mahoney, J. L., Weissberg, R. P., Greenberg, M. T., Dusenbury, L., Jagers, R. J., Niemi, K., Schlinger, M., Schlund, J., Shriver, T. P., VanAusdal, K., & Yoder, N. (2021) Systemic social and emotional learning: Promoting educational success for all preschool to high school students. *American Psychologist, 76*(7), 1128–1142. https://doi.org/10.1037/amp0000701

Manatos, M. J., Sarrico, C. S., & Rosa, M. J. (2017). The integration of quality management in higher education institutions: A systematic literature review. *Total Quality Management & Business Excellence, 28*(1–2), 159–175. https://doi.org/10.1080/14783363.2015.1050180

Massachusetts Consortium for Innovative Education Assessments (MCIEA). (n.d.). School quality measures framework. https://mciea-dashboard.herokuapp.com/welcome

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook.* Sage Publications, Inc.

Mitra, D. (2018). Student voice in secondary schools: The possibility for deeper change. *Journal of Educational Administration, 56*(5), 473–487. https://doi.org/10.1108/JEA-01-2018-0007

Morinaj, J., Scharf, J., Grecu, A., Hadjar, A., Hascher, T., & Marcin, K. (2017). School alienation: A construct validation study. *Front Learning Research, 5*(2), 36–59. https://doi.org/10.14786/flr.v5i2.298

Muhammad Butt, M., & Cyril de Run, E. (2010). Private healthcare quality: Applying a SERVQUAL model. *International Journal of Health Care Quality Assurance, 23*(7), 658–673. https://doi.org/10.1108/09526861011071580

Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin, 105*(3), 430–445. https://doi.org/10.1037/0033-2909.105.3.430

Newchurch, A. (2017). The impact of parental involvement on student success: School and family partnership from the perspective of parents and teachers [Doctoral dissertation, Kennesaw State University].

Newman, K. (2001). Interrogating SERVQUAL: A critical assessment of service quality measurement in a high street retail bank. *International Journal of Bank Marketing, 19*(3), 126–139. https://doi.org/10.1108/02652320110388559

Nikel, J., & Lowe, J. (2010). Talking of fabric: a multi-dimensional model of quality in education. *Compare: A Journal of Comparative and International Education, 40*(5), 589–605. https://doi.org/10.1080/03057920902909477

Odendaal, R. M., & Plessis, E. C. Du. (2018). The external school consultant as a proposed agent for school improvement. *The Independent Journal of Teaching and Learning, 13*(2), 82–97.

Parasuraman, A., Zeithaml, V. A. & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing, 64*(1), 12–40.

Pauley, J. A., & Pauley, J. F. (2009). *Communication: The key to effective leadership.* Milwaukee, WI: ASQ Quality Press.

Pedhazur, E. J., & Schmelkin, L. P. (2013). *Measurement, design, and analysis.* Psychology Press. https://doi.org/10.4324/9780203726389

Pfeffer, N., & Coote, A. (1991). *Is quality good for you?: A critical review of quality assurance in welfare services.* Institute for Public Policy Research.

Prichard Committee (2013). Prichard Committee for Academic Excellence. https://www.gatesfoundation.org/about/committed-grants/2013/07/opp1091659

Pulis, A. (2018). Mixed methods research on the role of pupils as assessors in quality assurance of schools in Malta. *Management in Education, 32*(1), 40–47. https://doi.org/10.1177/0892020617748152

Ramrathan, L. (2021). School curriculum in South Africa in the COVID-19 context: An opportunity for education for relevance. *PROSPECTS, 51*(1–3), 383–392. https://doi.org/10.1007/s11125-020-09490-1

Ramseook-Munhurrun, P. & Nundlall, P. (2013). Service quality measurement for secondary school setting. *Quality Assurance in Education, 21*(4), 387–401. https://doi.org/10.1108/QAE-05-2011-0025

Rayner, S. M. (2017), Admissions policies and risks to equity and educational inclusion in the context of school reform in England. *Management in Education, 31*(1), 27–32. https://doi.org/10.1177/0892020616687699

Razinkina, E., Pankova, L., Trostinskaya, I., Pozdeeva, E.,Evseeva, L., & Tanova, A. (2018). Student satisfaction as an element of education quality monitoring in innovative higher education institution. *E3S Web of Conferences, 33,* 03043. https://doi.org/10.1051/e3sconf/20183303043

Reimers, F., & Tiburcio, L. (1993). *Education, adjustment and reconstruction: Options for change*. UNESCO.

Reimers, F. M. (2020). What is global education and why does it matter? In *Educating students to improve the world* (pp. 25–29). Singapore: Springer. https://doi.org/10.1007/978-981-15-3887-2_2

Rosenkvist, M. A. (2010). Using student test results for accountability and improvement: A literature review. OECD Education Working Papers No. 54, OECD Publishing. https://doi.org/10.1787/5km4htwzbv30-en

Rust, C. (2002). The impact of assessment on student learning. *Active Learning in Higher Education, 3*(2), 145–158. https://doi.org/10.1177/1469787402003002004

Sahney, S., Banwet, D. K., & Karunes, S. (2004). Conceptualizing total quality management in higher education. *The TQM Magazine, 16*(2), 145–159. https://doi.org/10.1108/09544780410523044

Sakthivel, P.B., Rajendran, G., & Raju, R. (2005). TQM implementation and students' satisfaction of academic performance. *The TQM Magazine, 17*(6), 573–589. https://doi.org/10.1108/09544780510627660

Sallis, E. (2014). *Total quality management in education* (3rd ed.). Routledge. https://doi.org/10.4324/9780203417010

Sass, T. R., Semykina, A., & Harris, D. N. (2014). Value-added models and the measurement of teacher productivity. *Economics of Education Review, 38*, 9–23. https://doi.org/10.1016/j.econedurev.2013.10.003

Sayed, Y., & Ahmed, R. (2011). Education quality in post-apartheid South African policy: Balancing equity, diversity, rights and participation. *Comparative Education. 47*(1), 103–118. https://doi.org/10.1080/03050068.2011.541680

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of Structural Equation Models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*(2), 23–74.

Sfakianaki, E. (2019). A measurement instrument for implementing total quality management in Greek primary and secondary education. *International Journal of Educational Management, 33*(5), 1065–1081. https://doi.org/10.1108/IJEM-08-2018-0245

Shah, M., Nair, S., & Wilson, M. (2011). Quality assurance in Australian higher education: Historical and future development. *Asia Pacific Education Review, 12*(3), 475–483. https://doi.org/10.1007/s12564-011-9152-2

Smeby, J., & Stensaker, B. (1999). National quality assessment systems in the Nordic countries: Developing a balance between external and internal needs? *Higher Education Policy, 12*(1), 3–14. https://doi.org/10.1016/S0952-8733(98)00027-0

Smith, P. (1996). Tools for measuring quality improvement. *Management in Education, 10*(2), 21–23. https://doi.org/10.1177/089202069601000214

Stabback, P. (2016). *What makes a quality curriculum? Current and Critical Issues in Curriculum and Learning, 2.* UNESCO.

Stensaker, B., Langfeldt, L., Harvey, L., Huisman, J., & Westerheijden, D. (2011). An in-depth study on the impact of external quality assurance. *Assessment & Evaluation in Higher Education, 36*(4), 465–478. https://doi.org/10.1080/02602930903432074

Stepanova, N., Sannikova, L., Levshina, N., Yurevich, S. N., & Chernobrovkin, V. A. (2017). Parental evaluation of pre school education quality: Is it a problem or an opportunity? *Man in India, 97*(5), 171–185.

Swick, K. J., & Bailey, L. B. (2004). Communicating effectively with parents and families who are homeless. *Early Childhood Education Journal, 32*(3), 211–215. https://doi.org/10.1023/B:ECEJ.0000048975.59024.c4

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in Science education. *Research in Science Education, 48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tan, K. C., & Kek, S. W. (2004). Service quality in Higher Education using an enhanced SERVQUAL approach. *Quality in Higher Education, 10*(1), 17–24. https://doi.org/10.1080/1353832242000195032

Terrell, I., Terrell, K., Clinton, B., & Sheraton, K. (1996). Quality middle management: A d-i-y quiz on criteria. *Management in Education, 10*(1), 14–16. https://doi.org/10.1177/089202069601000107

Tikly, L., & Barrett, A. M. (2011). Social justice, capabilities and the quality of education in low income countries. *International Journal of Educational Development, 31*(1), 3–14. https://doi.org/10.1016/j.ijedudev.2010.06.001

Tran, T. Q., Pham, H. H., Vo, H. T., Luu, H. T., & Nguyen, H. M. (2019). Local governance, education and occupation-education mismatch: Heterogeneous effects on wages in a lower middle income economy. *International Journal of Educational Development, 71*, 102101. https://doi.org/10.1016/j.ijedudev.2019.102101

Tsuda, Y. (1995). TQM in education. In *Total quality management* (pp. 436–441). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-011-0539-2_76

Unterhalter, E. (2019). The many meanings of quality education: Politics of targets and indicators in SDG 4. *Global Policy, 10*(S1), 39–51. https://doi.org/10.1111/1758-5899.12591

Van Damme, D. (2011). VIII. Standards and indicators in institutional and programme accreditation in higher education: A conceptual framework and a proposal. https://www.semanticscholar.org/paper/VIII.-Standards-and-Indicators-in-Institutional-and-Damme/dcd2c457188634ea1993c4abfb1d4f81c1b5d5c5

Vietnam Ministry of Education and Training. (2006). Thông tư số 43/2006/TT-BGDĐT: Hướng dẫn thanh tra toàn diện nhà trường, cơ sở giáo dục khác và thanh tra hoạt động sư phạm của nhà giáo. [Circular number 43/TT-BGDDT: Instructions on the comprehensive inspection of schools, other educational institutions]. [In Vietnamese]

Vietnam Ministry of Education and Training. (2017). Quyết định số 2161/QĐ-BGDĐT: Ban hành kế hoạch thực hiện mục tiêu phát triển bền vững lĩnh vực giáo dục và đào tạo đến năm 2025 và định hướng đến năm 2030 [Decision Number 2161/QD-BGDDT: Issuing the plan to implement sustainable development goals in the field of education and training until 2025 and orientation towards 2030]. [In Vietnamese]

Vietnam Ministry of Education and Training. (2018a). *Chương trình giáo dục phổ thông mới* [New general education curriculum]. Ha Noi: Ministry of Education and Training.

Vietnam Ministry of Education and Training. (2018b). Thông tư số 18/2018/TT-BGDĐT: Quy định về kiểm định chất lượng giáo dục và công nhận đạt chuẩn quốc gia đối với trường trung học cơ sở, trường THPT và trường phổ thông có nhiều cấp học. [Circular Number. 18/2018/TT-BGDDT: Regulations on the accreditation of educational quality and recognition of national standards for secondary schools, high schools, and multi-level schools.] [In Vietnamese]

Vuong, Q. -H. (2018). The (ir)rational consideration of the cost of science in transition economies. *Nature Human Behaviour, 2*, 5. https://doi.org/10.1038/s41562-017-0281-4

Vuong, Q. -H. (2020). Reform retractions to make them more transparent. *Nature, 582*(7811), 149–149. https://doi.org/10.1038/d41586-020-01694-x

Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. *Sociological Methodology, 8*, 84–136. https://doi.org/10.2307/270754

William, D. (2010). Standardized testing and school accountability. *Educational Psychologist, 45*(2), 107–122. https://doi.org/10.1080/00461521003703060

Williams, P. J. (2011). STEM education: Proceed with caution. *Design and Technology Education, 16*(1), 26–35. https://doi.org/10.24377/DTEIJ.article1748

Wöbmann, L., Lüdemann. E., Schütz. G., & West, M. R. (2007). School accountability, autonomy, choice, and the level of student achievement. OECD Education Working Papers No. 13. https://doi.org/10.1787/246402531617

Wong, M. N. C, & Li, H. (2010). From external inspection to self-evaluation: A study of quality assurance in Hong Kong kindergartens. *Early Education & Development, 21*(2), 205–233. https://doi.org/10.1080/10409281003638725

Yocke, M. (1997). Can quality be total? A letter from the University of Erewhon. *Management in Education, 11*(3), 26–27. https://doi.org/10.1177/089202069701100313

Yousapronpaiboon, K. (2014). SERVQUAL: Measuring higher education service quality in Thailand. *Procedia – Social and Behavioral Sciences, 116*, 1088–1095. https://doi.org/10.1016/j.sbspro.2014.01.350

Zedan, R. (2010). New dimensions in the classroom climate. *Learning Environments Research, 13*(1), 75–88. https://doi.org/10.1007/s10984-009-9068-5

APPENDICES

Appendix A: KMO and Barlett's Test

		Expectation	Perception
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.833	.764
Bartlett's Test of Sphericity	Approx. Chi-Square	25,957.177	12,748.644
	df	171	171
	Sig.	0.000	0.000

Appendix B: Pattern Matrix[a] – part 1

	Component					
	1	2	3	4	5	6
Exp_Cli_Safe	.882					
Exp_Cli_Clean	.788					
Exp_Cli_Love	.776					
Exp_Cli_Res	.665					.358
Exp_Teach_Aid		.934				
Exp_Teach_Achi		.915				
Exp_Teach_Ins		.886				
Exp_Lead_Cap			.962			
Exp_Lead_Mana			.940			
Exp_Lead_Goal			.796			
Exp_Res_Cam				.913		
Exp_Res_Human				.852		
Exp_Res_Inf				.633		
Exp_Cur_Vari					.901	
Exp_Cur_Update					.843	
Exp_Ass_Cri	.468				.527	
Exp_Com_Ins						.814
Exp_Com_Exp	-.316					.765
Exp_Com_Par	.506					.660

Extraction Method: Principal Component Analysis.
Rotation Method: Promax with Kaiser Normalisation.
a. Rotation converged in 7 iterations.

Appendix C: Pattern Matrix[a] – part 2

	Component					
	1	2	3	4	5	6
Per_Cli_Res	.843					
Per_Cli_Safe	.830					
Per_Cli_Clean	.785					
Per_Cli_Love	.692					
Per_Teach_Ins		.904				
Per_Teach_Aid		.887				

	Component					
	1	2	3	4	5	6
Per_Teach_Achi		.864				
Per_Res_Cam			.856			
Per_Res_Human			.806			
Per_Res_Inf			.794			
Per_Com_Exp				.869		
Per_Com_Ins				.803		
Per_Com_Par				.721		
Per_Cur_Update					.857	
Per_Ass_Cri					.812	
Per_Cur_Vari					.699	
Per_Lead_Mana						.772
Per_Lead_Goal						.758
Per_Lead_Cap						.755

Extraction Method: Principal Component Analysis.
Rotation Method: Promax with Kaiser Normalisation.
a. Rotation converged in 5 iterations.

Appendix D: Paired samples statistics

		Mean	*N*	SD	Std. Error Mean
Pair 1	Per_Lead_Goal	3.56	2,239	1.075	.023
	Exp_Lead_Goal	4.04	2,239	.802	.017
Pair 2	Per_Lead_Cap	3.72	2,239	.951	.020
	Exp_Lead_Cap	4.04	2,239	.793	.017
Pair 3	Per_Lead_Mana	3.82	2,239	.946	.020
	Exp_Lead_Mana	4.04	2,239	.777	.016
Pair 4	Per_Teach_Ins	3.49	2,239	.993	.021
	Exp_Teach_Ins	4.04	2,239	.817	.017
Pair 5	Per_Teach_Aid	3.34	2,239	.885	.019
	Exp_Teach_Aid	4.02	2,239	.805	.017
Pair 6	Per_Teach_Achi	3.77	2,239	.790	.017
	Exp_Teach_Achi	4.00	2239	.793	.017
Pair 7	Per_Cur_Vari	3.46	2,239	.983	.021
	Exp_Cur_Vari	4.09	2,239	.752	.016

		Mean	*N*	SD	Std. Error Mean
Pair 8	Per_Cur_Update	3.45	2,239	1.034	.022
	Exp_Cur_Update	4.07	2,239	.796	.017
Pair 9	Per_Ass_Cri	3.55	2,239	.828	.017
	Exp_Ass_Cri	4.09	2,239	.790	.017
Pair 10	Per_Cli_Clean	3.24	2,239	.848	.018
	Exp_Cli_Clean	4.11	2,239	.798	.017
Pair 11	Per_Cli_Safe	3.38	2,239	.805	.017
	Exp_Cli_Safe	3.96	2,239	.742	.016
Pair 12	Per_Cli_Res	3.29	2,239	.861	.018
	Exp_Cli_Res	4.19	2,239	.835	.018
Pair 13	Per_Cli_Love	3.19	2,239	.722	.015
	Exp_Cli_Love	3.79	2,239	.623	.013
Pair 14	Per_Com_Par	3.37	2,239	.899	.019
	Exp_Com_Par	4.15	2,239	.758	.016
Pair 15	Per_Com_Ins	3.20	2,239	.795	.017
	Exp_Com_Ins	3.97	2,239	.750	.016
Pair 16	Per_Com_Exp	3.02	2,239	.728	.015
	Exp_Com_Exp	4.06	2,239	.745	.016
Pair 17	Per_Res_Human	3.46	2,239	.751	.016
	Exp_Res_Human	3.81	2,239	.730	.015
Pair 18	Per_Res_Inf	3.77	2,239	.810	.017
	Exp_Res_Inf	4.09	2,239	.693	.015
Pair 19	Per_Res_Cam	3.87	2,239	.738	.016
	Exp_Res_Cam	4.17	2,239	.765	.016

Appendix E: Paired samples correlation

		N	Correlation	Sig.
Pair 1	Per_Lead_Goal & Exp_Lead_Goal	2,239	.566	.000
Pair 2	Per_Lead_Cap & Exp_Lead_Cap	2,239	.362	.000
Pair 3	Per_Lead_Mana & Exp_Lead_Mana	2,239	.587	.000
Pair 4	Per_Teach_Ins & Exp_Teach_Ins	2,239	.426	.000
Pair 5	Per_Teach_Aid & Exp_Teach_Aid	2,239	.266	.000
Pair 6	Per_Teach_Achi & Exp_Teach_Achi	2,239	.301	.000

		N	Correlation	Sig.
Pair 7	Per_Cur_Vari & Exp_Cur_Vari	2,239	.662	.000
Pair 8	Per_Cur_Update & Exp_Cur_Update	2,239	.676	.000
Pair 9	Per_Ass_Cri & Exp_Ass_Cri	2,239	.205	.000
Pair 10	Per_Cli_Clean & Exp_Cli_Clean	2,239	.473	.000
Pair 11	Per_Cli_Safe & Exp_Cli_Safe	2,239	.649	.000
Pair 12	Per_Cli_Res & Exp_Cli_Res	2,239	.381	.000
Pair 13	Per_Cli_Love & Exp_Cli_Love	2,239	.461	.000
Pair 14	Per_Com_Par & Exp_Com_Par	2,239	.530	.000
Pair 15	Per_Com_Ins & Exp_Com_Ins	2,239	.560	.000
Pair 16	Per_Com_Exp & Exp_Com_Exp	2,239	.559	.000
Pair 17	Per_Res_Human & Exp_Res_Human	2,239	.453	.000
Pair 18	Per_Res_Inf & Exp_Res_Inf	2,239	.631	.000
Pair 19	Per_Res_Cam & Exp_Res_Cam	2,239	.601	.000

Appendix F: Paired samples test

		Paired differences					*t*	df	Sig. (2-tailed)
		Mean gap diff.	SD	Std. error mean	95% confidence interval of the difference				
					Upper	Lower			
Pair 1	Per_Lead_ Goal - Exp_ Lead_Goal	-.483	.908	.019	-.521	-.446	-25.197	2,238	.000
Pair 2	Per_Lead_ Cap - Exp_ Lead_Cap	-.327	.993	.021	-.369	-.286	-15.599	2,238	.000
Pair 3	Per_Lead_ Mana - Exp_ Lead_Mana	-.216	.797	.017	-.249	-.183	-12.835	2,238	.000
Pair 4	Per_Teach_ Ins - Exp_ Teach_Ins	-.546	.981	.021	-.587	-.506	-26.352	2,238	.000
Pair 5	Per_Teach_ Aid - Exp_ Teach_Aid	-.680	1.026	.022	-.722	-.637	-31.353	2,238	.000
Pair 6	Per_Teach_ Achi - Exp_ Teach_Achi	-.234	.935	.020	-.273	-.196	-11.862	2,238	.000

		Paired differences					*t*	df	Sig. (2-tailed)
		Mean gap diff.	SD	Std. error mean	95% confidence interval of the difference				
					Upper	Lower			
Pair 7	Per_Cur_Vari - Exp_Cur_ Vari	-.637	.743	.016	-.668	-.606	-40.555	2,238	.000
Pair 8	Per_Cur_ Update - Exp_Cur_ Update	-.616	.768	.016	-.648	-.585	-37.983	2,238	.000
Pair 9	Per_Ass_Cri - Exp_Ass_Cri	-.540	1.020	.022	-.582	-.498	-25.040	2,238	.000
Pair 10	Per_Cli_ Clean - Exp_ Cli_Clean	-.870	.846	.018	-.906	-.835	-48.677	2,238	.000
Pair 11	Per_Cli_Safe - Exp_Cli_ Safe	-.588	.650	.014	-.615	-.561	-42.802	2,238	.000
Pair 12	Per_Cli_Res - Exp_Cli_Res	-.899	.943	.020	-.938	-.860	-45.070	2,238	.000
Pair 13	Per_Cli_Love - Exp_Cli_ Love	-.599	.704	.015	-.629	-.570	-40.314	2,238	.000
Pair 14	Per_Com_Par - Exp_Com_ Par	-.784	.812	.017	-.818	-.751	-45.684	2,238	.000
Pair 15	Per_Com_Ins - Exp_Com_ Ins	-.767	.726	.015	-.797	-.737	-50.007	2,238	.000
Pair 16	Per_Com_ Exp - Exp_ Com_Exp	-1.034	.692	.015	-1.063	-1.006	-70.775	2,238	.000
Pair 17	Per_Res_Inf - Exp_Res_Inf	-.324	.654	.014	-.351	-.297	-23.440	2,238	.000
Pair 18	Per_Res_Cam - Exp_Res_ Cam	-.301	.672	.014	-.329	-.273	-21.192	2,238	.000
Pair 19	Per_Res_ Human - Exp_Res_ Human	-.346	.775	.016	-.378	-.314	-21.145	2,238	.000