A Thesis for the Degree of Ph.D. in Engineering

A model of total quality management and student score analysis in school education

February 2023

Graduate School of Science and Technology
Keio University

Alauddin, Nursuhana Binti

ABSTRACT

Educational institutions, including schools, prioritize educational quality to provide students with the finest possible learning experiences. One of the useful tools to improve and enhance quality in education is the implementation of total quality management (TQM). Frameworks for educational systems have been established as part of continual improvement via quality standards and frameworks to ensure that educational institutions perform their goals. In this research, a TQM model based on Deming criteria for school education is developed to ensure organizational strength in school management in the present and attaining sustainable excellence in the future. A qualitative approach is carried out to validate the model at an international school in Japan. Furthermore, as education is mainly about student learning, students are the center of education. The development of student monitoring systems enables the observation of the growth of individual students and learning progress on a longitudinal basis. The systems are designed to manage various parameters to improve educational quality and encourage data-based decision-making through the provision of data for goal setting. Owing to the progression of information technology systems, not only semester examination data, but also daily assessment data are becoming available in the school's centralized monitoring system. Therefore, it is made possible to utilize such data for immediate remedies and hints of a good practice in the same academic year based on student ability, subjects, and various assessments. In this research, two methods are developed for the detection of unusual scores for individual students according to the type of data and purpose of detection; 1) analysis based on daily assessment data through a linear ANOVA model to detect unusual scores in a daily assessment, and 2) analysis based on past scores, current daily efforts, and current score trends by a combination of ANOVA, principal component analysis, and multiple regression analysis to detect unexpected score in an examination. The unusual or unexpected score is detected simultaneously for different subjects for individual students. Dataset is obtained from an international school in Japan for model validation.

TABLE OF CONTENTS

CHAPTER 1: Introduction

1.1 Research Background

Enhancing quality is perhaps the most essential challenge confronting any organization. Despite its importance, many people regard quality as a subjective term. Because quality is a relative notion to different individuals, it may be described in many viewpoints and aspects. In education as well, quality is difficult to define and measure, but its absence is all too obvious. In the 2030 Agenda of Sustainable Development Goal (SDG), SDG4 aims to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all (UNESCO, 2017). Quality in educational systems involves more than merely distinguishing between the exceptional and the ordinary, or between success and failure. Educational institutions cope with the competitive world by providing high-quality education and continuously improving their processes (Venkatraman, 2007). Arcaro (1995) opines that quality objectives for education should reflect areas of community responsibility and citizenship. The educational quality plans should contain responsiveness to community needs and processes to develop and maintain public trust. Sallis (2002) mentions that four quality imperatives drive educational institutions for quality improvements. The complexity of education and the importance of values in education have made the motives for taking a quality stance more complicated and diverse. Those are imperatives based on moral, professional, competitive, and accountability.

There are several difficulties facing by schools in implementing quality management. The successful implementation of quality management in education depends on full cooperation among stakeholders. The school culture should include all stakeholders in defining the organization goals. Nevertheless, there is a wide range of customers with conflicting expectations in education. Researchers typically draw attention to the complexity of the customer definition in education, in contrast to the straightforward customer interpretation in industries (Koch, 2003; Mehta and Degi, 2019). In addition to the ‘main customers’, which are the parents and students, there are many more stakeholders. Contrarily, if too many different types of stakeholders are considered, it may be difficult to focus on the ‘real customer’, which are the students, and to gauge the success of quality efforts. Besides, lack of understanding of quality management's principles and related tools to carry out improvement activities is also a huge obstacle in education (Bouranta *et al*., 2020). Arcaro (1995) concurs that many of educational professionals are lack of knowledge on how to cope with new requirements expected to them and this may generate fear about the unknown and discomfort with doing things differently. Ozberk *et al.* (2019) agrees that in some cases, school administrators are ignorant about quality management and none of the teachers and school administrators have previously heard about the concept of quality tools. Without a clear understanding of the institutions' missions and ability to implement TQM activities to achieve goals, there will unavoidably be a lack of unity of purpose and the potential for quality work to be neglected. Thus, it is impossible to move forward with process management and evaluating the effectiveness of the offered services toward attaining school excellence.

Generally, education is about educating human beings; thus, students should be in the center of learning. Students vary in the ability and process of learning, so understanding diverse learning needs is important in educational institutions. The students enter the classroom with different backgrounds, skill sets, and educational needs. According to Felder and Brent (2005), three aspects of student diversity—differences in learning styles, learning strategies, and intellectual development levels—have a substantial impact on teaching and learning. The more thoroughly teachers understand the differences, the better chance they have of meeting the diverse learning needs of all students. Figuring out how to properly meet each student's learning needs is challenging. Teachers often find it difficult to respond to student queries and adjust their pedagogies to each student's level due to the wide range of subjects, assessments, and curriculum standards in one academic year. Kwan (1996) highlights that because teaching and learning are collaborative processes involving teachers and students, they cannot be preset in a step-by-step manner like an assembly line. Additionally, as both parties are humans, their behavior is influenced by a variety of goals and motivations, emotional swings, and personal interpersonal styles. As a result, there cannot be any specification as regards the standardized codes of practice in teaching and learning. When it comes to inputs, educational settings are subject to enormous variability. In order to produce standard outputs, it is challenging to establish process assurance and daily management in the teaching and learning process.

Several approaches can be taken to improve educational quality. One of them is discipline-based educational research (DBER). This approach is to investigate learning and teaching in a discipline's priorities, worldview, knowledge, and practices. As the interdisciplinary nature of this approach involves social science, the faculty members of science, technology, engineering, and mathematics may find it challenging to evaluate the impact or quality of discipline-based educational research scholarship (Dolan *et al.*, 2018). Another way to improve the quality of education is through educational effectiveness research (EER), which attempts to establish and test theories that explain why and how some schools and teachers are more effective than others. Scheerens (2016) explains educational effectiveness research and its improvement, as it starts from an integrated multi-level model that contains system level, school, level, and instructional conditions. It suggests a classification of school improvement strategies and situations for system-level educational improvement. Also, human capital theory (HCT), which is a framework that examines the relationships between education, economic growth, and social well-being may be one of the methods to improve quality. In this theory, the more and better education individuals possess, the better their returns in financial rewards and the better the national economy flourishes. It promotes education to a key instrumental role in boosting economic growth (Gillies, 2015).

Another potential approach that may improve the quality of education is total quality management (TQM). TQM is a philosophy, concept, and a collection of tools that may be used in the management of educational institutions. TQM concept is originally created for the manufacturing sectors; thus, it is debatable whether it can be adopted in education. Sallis (2002) opines that it is more appropriate to view education as a service industry rather than as a production process. Once this understanding is established, the institution needs to clarify their services and the standards to which they will be delivered. Beaver (1994) states that traditionally, defining quality has been left up to individual educators. Under TQM, quality is achieved by meeting the customers' expectations, as customer plays a major role in defining quality. To avoid misinterpretation of TQM philosophies and practices, Sallis argues that TQM must avoid being nothing but jargon and hype. Sharing the missions of the schools can help close communication gaps between top management and staffs, and ensure that everyone is aware about their organization's direction and how to make it happen to achieve school excellence and long-term sustainability. Sfakianaki (2019) states that researchers have focused on the TQM implementation more in higher educational institutions and its implementation in primary and secondary education remains scant. This is despite the current surge in interest in primary and secondary education as a sector that provides more widespread social gains, while also serving as the most significant locus for the establishment of fundamental values and the basis for educational super-systems. Therefore, the TQM model which is appropriate for school education is chosen as the approach to improve the quality of education.

A suitable theory of education should treat the teaching and learning as important to the institution's mission, and it should bring methods for improvement to bear on these activities (Tribus, 2005). As teaching and learning is a core process in education and happens on daily basis, it must be managed methodically through a monitoring system that provides fast feedback to school management. The development of a student monitoring system enables to the observation of the growth of individual students and learning progress on a longitudinal basis. The systems are designed to manage various parameters to improve educational quality and encourage data-based decision-making through the provision of data for goal setting (Vlug, 1997). There are two types of monitoring systems: professional monitoring systems and accountability systems. The primary goal of professional monitoring systems, such as assessment systems, student monitoring systems, and school self-evaluation systems, is to improve schools, whereas the accountability system is to hold schools responsible as publicly supported institutions. Monitoring systems can employ a variety of data analysis methodologies, ranging from complex analysis that yields statistically correct and reliable findings, to simple analysis that lead to user-friendly results (Schildkamp and Archer, 2017).

Making good use of data may lead to improvement in the quality of education. Data-driven decision-making in education typically refers to teachers, principals, and administrators systematically analyzing data to guide a range of decisions to help improve the success of students and schools. Data can be defined as quantitative as well as qualitative information that is systematically collected and organized to represent some aspect of schooling. School management and teachers might utilize multiple types of data, including input data, such as demographics of the student population; process data, for example, data on financial operations; outcome data, such as student test scores; and satisfaction data, for instance, opinions from teachers, students or parents (Ikemoto and Marsh, 2007). It is suggested that a framework of the process that multiple forms of data are first turned into information via analysis and then combined with stakeholder understanding to create actionable knowledge. To produce information, raw data are coupled with a knowledge of the issue by an analysis process. Data users may turn information into actionable knowledge by prioritizing information and weighing the merits of potential solutions. Once a choice has been implemented, additional data may be acquired to evaluate the success of those actions. As such, it is resulting in a continuous cycle of data collection and organization to support decision-making.

Utilizing student score data may lead to a huge improvement in the teaching and learning process. Student scores are gathered through standardized examinations and assessments, which come in several formats and are designed to serve a variety of reasons. They include machine-scored, multiple-choice examinations, assessments demanding short or long answers, and some sort of hands-on performance evaluation assignments. Accountability demands have resulted in significant increases in the number of testing and assessments in schools, as well as in the importance that is attached to the results (Miller *et al*., 2009). A betterment in student performance is one of the most important missions in school education. In comparison to average-ability students, the low-ability students might find it difficult to keep up with the lessons and remedial action taken by teachers to improve their learning experiences may not be sufficient. On the other hands, high-ability students are often experienced boredom, leading to a loss of motivation, which in turn can lead to underachievement while learning in courses aimed at average students.

Teaching and learning process in schools happens every day with a standard curriculum; thus, it shall be managed through daily management practices. In daily management, the detection of abnormality is essential in order to stabilize the process to achieve target (JSQC, 2014). Abnormality, or can be called as unusual or unexpected point, represents data where the process is out of normal condition. In this research, the utilization of data based on student scores is chosen to enhance the process assurance, particularly in management of teaching and learning. Combining the concept of detection of unusual or unexpected condition in daily activities considering students' diverse abilities in various type of assessments may help to enhance the process assurance in education.

1.2 Research questions

There are three research questions to be answered through this research as below:

1) How to develop a TQM model for school education to realize school objectives and enhance organizational ability in the present and future?

2) How to develop a method to detect the unusual condition of individual students in daily activities, based on daily assessment data to enhance daily management practices in schools?

3) How to develop a method to detect the unexpected condition of individual students in an examination, based on a combination of past examination data and current daily assessment data to enhance daily management practices in schools?

1.3 Outlines of the dissertation

This dissertation is organized into five chapters, in addition to References, Acknowledgements, and Appendices. This chapter provides an introduction to the research describing the background, research questions, and the outline of the research.

Chapter 2 is dedicated to answering research questions 1). It describes the development of a TQM model for school education to realize school objectives and obtain organizational ability in the present and attain sustainable excellence in the future. A qualitative research approach is utilized for model validation through a case study.

Chapter 3 is aimed to answer research question 2), introducing research about the detection of unusual scores in a daily assessment for individual students. The detection is described through a linear model based on daily assessment data. A dataset from an international school is obtained to validate the model.

Chapter 4 is aimed to answer research question 3) for unexpected condition detection. It describes research about the detection of unexpected scores in an examination for individual students. The detection is described through a linear model based on data of past scores, current daily efforts, and current score trends. A different dataset from Chapter 3 is obtained from an international school to validate the model.

Chapter 5 is a concluding remark to synergize what has been achieved in the above research as a total.

CHAPTER 2: TQM Model based on Deming prize for schools

2.1 Introduction

Over the past several decades, total quality management (TQM) has been widely adopted in the industrial (Chapman and Al-Khawaldeh, 2002) and service sectors (Khalaf and Salem, 2018), and has been practiced successfully by countless organizations worldwide. Educational institutions also have begun to adopt TQM, with positive results (Sallis, 2002; Töremen *et al.*, 2009). Appropriate models for TQM implementation can clarify the activities that may contribute to organizational improvement, allowing an organization to clearly discern its strengths and weaknesses. This culminates in improvement activities whose progress can be actively monitored (van der Wiele *et al.*, 1995).

There are a number of quality awards recognizing successful business applications of TQM. These include the Malcolm Baldrige National Quality Award (MBNQA) and the European Foundation for Quality Management (EFQM) award. The frameworks of these awards share similar principles, and their commonalities far outweigh their differences. ISO standard and MBNQA models are among the most widely used quality models in a variety of sectors. These models are basically prescriptive, with guidelines regarding principles and systems that provide an organization with a clear path to be followed. Accordingly, multiple organizations may achieve similar results with the same standards. These standards tend to encourage gradual improvement and the assessment of organizations based on their existing performance and results.

Frameworks for educational systems have been established as part of continual improvement in organizational learning (Murgatroyd and Morgan, 1993; Dalin, 2005). These frameworks have been built via quality standards and awards to ensure that educational institutions perform their goals. Some examples are the European Education Quality Benchmark System, the Baldrige Excellence Framework (Education), the EFQM-Hamdan Education Model, and ISO 21001:2018 (Snyder, 2007; NIST, 2015; EFQM, 2021; ISO, 2018).

The Deming Prize is another prominent quality award; it recognizes organizations that have obtained sustainable business success through TQM implementation. It provides a generic framework that emphasizes customer-oriented business objectives and strategies that drive

TQM activities and are reflected in organizational results and impact. The Deming framework focuses on organizational ability, specifically both current operational performance and future growth. Its criteria aim at achieving the organization's customer-oriented objectives through three main categories: A (direction), B (activities), and C (effects). The Deming framework is somewhat ambiguous; hence, it allows organizations more freedom in determining suitable TQM activities. Therefore, without proper guidelines, an organization may face serious challenges in implementing this framework. Despite this issue, or perhaps because of it, little research has been conducted with a focus on the Deming Prize (Thandapani *et al.*, 2012) as compared to other quality awards, models and management systems. Although the Deming framework is a generic model and can be used by all sectors, the educational institutions may face difficulties to understand the criteria and terminologies.

The aim of this chapter is to develop a TQM model for education based on Deming criteria from the Deming Prize. Use of Deming criteria in schools can serve as a catalyst for ensuring organizational strength in school management in the present and attaining sustainable excellence in the future. It enables school management to link the effects of TQM implementation to school objectives and strategies, and offers systematic guidelines for executing effective TQM activities in the school. The model developed in this chapter can be used by schools to guide the adoption of TQM activities and further school excellence. The effectiveness of the model is assessed through a case study in which the model is applied at an international school.

2.2 Total quality management practices and models

2.2.1 Total quality management implementation in education

In the context of TQM implementation in education, Kwan (1996) points out that research appears to suggest that TQM is a unique solution for effective school management. In this regard, Sallis (2002) stresses that it is more appropriate to view education as a service industry rather than as a production process. Although some education specialists disagree with the importation of manufacturing analogies into the educational process, the implementation of TQM at the school level, particularly in primary schools, is gradually increasing (Töremen *et al.*, 2009). Sallis argues that if quality is about meeting customers' needs, it is essential to clarify whose needs it should be satisfying. To some educators, 'customer' has a distinctly commercial tone that is not applicable to education. They prefer to use client, stakeholder, and some of them reject all such language and would rather stay with pupil or student.

Research related to TQM implementation and the adoption of quality management models has been conducted by various educational institutions in different regions. Goldberg and Cole (2002) describes an exemplary school district that uses a quality management approach in terms of philosophy, tools, and methods, resulting in higher student performance. Sakthivel and Raju (2006) proposes a model that attempts to implement TQM in engineering education; it is built through an understanding of the requirements and needs of students as a prerequisite of enhancing the quality of their education. Johnson and Golomskiis (1999) asserts that leaders should establish unified goals in education and be responsible for ensuring that relevant parties are fully involved. Furthermore, continuous evaluation should be carried out through internal and external audits that enable analysis of quality indicators and non-conformance (Arribas Díaz and Martínez-Mediano, 2018). Murgatroyd and Morgan (1993) constructs a TQM framework that delivers tools from the context of schooling to school-based management. Dalin (2005) provides a comprehensive model for school improvement based on a theoretical foundation and strategies for implementing educational changes that address new twenty-first century realities.

One way to assess TQM implementation in schools is through assessment based on TQM models. There are some drawbacks associated with conducting model-based assessments, not least that the work is time-consuming and resource-demanding (Enquist *et al.*, 2015). Models for assessing educational activities are available, as demonstrated by the Baldrige Excellence Framework (Education), which introduced an education category in 1999. A similar model is utilized at some higher education institutions (Ruben *et al.*, 2007). Moreover, following the adoption of the EFQM framework, winners have been selected in the EFQM award school category since 2001. Anastasiadou *et al.* (2014) adopts a similar model to consider the needs of stakeholders and processes to improve educational performance.

2.2.2 Total quality management models

With respect to the MBNQA, according to the National Institute of Standards and Technology, interrelated criteria, core values, and scoring guidelines need to be integrated as a mechanism to achieve successful performance (NIST, 2015). While in the MBNQA, the emphasis is primarily on current organizational status, consideration of future plans is highlighted in the

Baldrige Excellence Framework (Education). Elsewhere, the EFQM has established the EFQM-Hamdan Education Model as a recognized framework to support the education sector (EFQM, 2021). Along with the quality management standard, ISO 9001:2015, it specifies the requirements of fundamental beliefs and values that serve as the basis for organizational quality management systems (ISO, 2015). Nakajo *et al*. (2015) state that ISO 9001 is not aimed at TQM, although the scope of its application has expanded to improve customer satisfaction. Another useful standard is ISO 9004:2018, which is helpful in advancing efforts based on human imperfections and changes in organizational conditions; however, such efforts should always evolve according to the circumstances of the organization (Nakajo *et al.*, 2019). ISO 21001:2018 is a specific ISO standard for educational institutions that focuses on their management systems and how they impact learners and other interested parties (ISO, 2018).

Given that the ISO and MBNQA provide a framework and specific requirements for each of their criteria, organizations can readily adopt these models as a means of achieving business excellence. The Baldrige Excellence Framework (Education) comprises seven key categories of organizational management (of which six are interrelated process categories and one is a results category): leadership, strategy, customers, measurement, analysis and knowledge management, workforce, operations, and results. The Baldrige Excellence Framework (Education) encourages organizations to choose tools, such as balanced scorecards, that are most effective for their needs. The criteria from this framework are prescriptive and according to descriptions of the submitted document to the assessors based on 'how' to carry out activities, such as 'how do you conduct your strategic planning?'. An example of descriptions of a criterion from the Baldrige Excellence Framework (Education) is included in Appendix 2. As a result, how organization thinks what should be done for their organization can be challenging for evaluation. ISO 9001:2015 separately defines customer focus and leadership, and specifies the ways in which leaders establish organizational objectives. With its management system for educational organizations, ISO 21001:2018 establishes eleven management principles: focus on learners and other beneficiaries, visionary leadership, engagement of people, process approach, improvement, evidence-based decisions, relationship management, social responsibility, accessibility and equity, ethical conduct in education, and data security and protection (ISO, 2018).

In regards to the Deming Prize, under certain conditions, any organization may apply to participate in the award regardless of their business type. The evaluation items for the Deming Prize consist of the following:

Category A

(I) Establishment of proactive customer-oriented business objectives and strategies
(II) Role of top management and its fulfilment

Category B

(III) Suitable utilization and implementation of TQM for the realization of business objectives and strategy

Category C

(IV) Organization has obtained effects on business objectives and strategies through suitable utilization and implementation of TQM
(V) Outstanding TQM activities and acquisition of organizational capabilities

Category B is further divided into seven criteria, from B1 to B7. The details of the evaluation items are applicable to the entire manufacturing and services sectors, and can be referenced at the web page of the Union of Japanese Scientists and Engineers (JUSE, 2021).

Vokurka *et al.* (2000) explains the Deming Prize from a broad perspective, relating it to the MBNQA. Their study condenses the eleven criteria of the Deming Prize's A, B, and C categories into ten elements. Kumar (2007) highlights that while the Deming Prize examines organizations using the evaluation criteria contained in the basic categories to assess the organization's unique activities and the role of its top management, the MBNQA assesses an organization based on the single additive dimension of its composite score. Although winning a quality award is not the main reason for TQM implementation, aiming to win allows an organization to measure success, use suitable diagnostic methods and identify opportunities for improvement (Dale *et al*., 2016). According to the JUSE (2021) document, there have been 258 Deming Prize winners since its inception (as of 2020). Currently, none of these winners have been from the educational sector.

The Deming Prize highlights customer-oriented business objectives and strategies that provide direction for an organization's TQM activities, as indicated in category A. The evaluation covers the organization's ability to promote sustainability in its forthcoming paths and to confront challenges to achieve business excellence. The Deming Prize links results and impact to customer-oriented business objectives and strategies through TQM implementation, as can be seen from the overall categorization. Alauddin and Yamada (2019) provides an overview of Deming criteria for the construction of a TQM conceptual framework suitable to education services from a process-based perspective.

This chapter proposes a comprehensive TQM model rooted in Deming criteria, adapted to the school context. The TQM approach considers the system of schooling as a whole system and integrates the interactions of all activities to assure quality in learning. By using a TQM model based on Deming criteria, schools can expect to attain effective organizational ability in education management, not only for the present but also for advancement towards future excellence. With this approach, schools are encouraged to formulate appropriate school objectives and strategies as one of the essential elements of success. These objectives should stimulate a balanced recognition of all stakeholder interests in order to gain support for their realization. Guided by the model, school management should be able to link the results and effects of TQM implementation to the school's objectives and strategies, as formulated in the initial stage. In addition, schools will come to understand how TQM activities can best be carried out through a systematic, process-based approach.

2.3 Total quality management model based on Deming criteria for school education

2.3.1 Basic idea and process of developing the total quality management model

Three basic ideas are reflected in the model development process. These ideas allow for the creation of a comprehensive model congruent with school education that can ensure organizational ability to assess and improve current performance and promote future growth:

(i) The proposed TQM model should encourage a high level of implementation that suits the school environment and can be benchmarked for others by using the Deming criteria's A, B and C categories structure to promote a sound organizational ability. In constructing the model, the A, B and C categories structure serves as the overall framework.

(ii) The proposed TQM model should identify relevant stakeholders through the provision of value in learning, skills development, and communication among the stakeholders. The stakeholders involved in school education are classified into groups by their importance in the education process. In the model and in this chapter, 'customer-oriented' is described instead as ‘society-oriented’ partly because of the perceived inappropriateness of using customer terminology and the complex nature of defining ‘customers’ in education.

(iii) For each of the A, B and C categories, each criterion is described in detail through separate sub-criteria to promote a clear understanding of each activity. The sub-criteria are described in a single statement for easy recognition.

To develop the comprehensive A, B and C criteria and sub-criteria, these four steps are followed:

Step 1: Rephrase each criterion to fit the school context

Each criterion and evaluation item are rephrased to fit the school context in order to simplify the subsequent steps of finding literature and identifying relevant factors. For example, one of the criteria for the Deming Prize refers to the ‘creation of new values based on an understanding of customer and social needs and innovation of technology and business model’. This is rephrased as the ‘creation of new educational values based on an understanding of societal needs’ and ‘education utilizing innovative technology’.

Step 2: Selection of literature related to each criterion

To identify the applicable literature for developing the appropriate sub-criteria, articles focusing on keywords related to each criterion are identified. Studies and articles on both general practice and the educational context are targeted.

Step 3: Identification of factors for each criterion

Required factors on how to carry out TQM activities related to the criteria are determined from the literature on both general and education-specific practice. These factors are used as sub-criteria to provide a systematic process-based approach to school activities.

Step 4: Finalize sub-criteria and rephrase into educational terminologies

The selected factors are finalized as sub-criteria based on their appropriateness for systematic practice suited to the school context. The sub-criteria related to key concerns of TQM activities such as being society-oriented, showing continuous improvement, aiming for total involvement, and so forth. For example, continuous improvement is included in sub-criteria of B3b, which emphasizes that activities carried out at schools should improve the quality of education and processes. It also states that all levels in the school organization should be involved in continuous improvement activities and that these activities should be embedded in school culture.

2.3.2 Detailed model description

2.3.2.1 Overall model framework

Based on the basic idea (i) presented in Section 2.3.1, this chapter applied the Deming criteria's A, B and C structure to education. Figure 2.1 and Table 2.1 illustrate the TQM model developed for schools. The A, B and C criteria are rephrased according to Step 1 above; the results are shown in Table 2.1. The model consists of the key elements for implementing TQM and guidance regarding how to best perform each activity comprehensively.

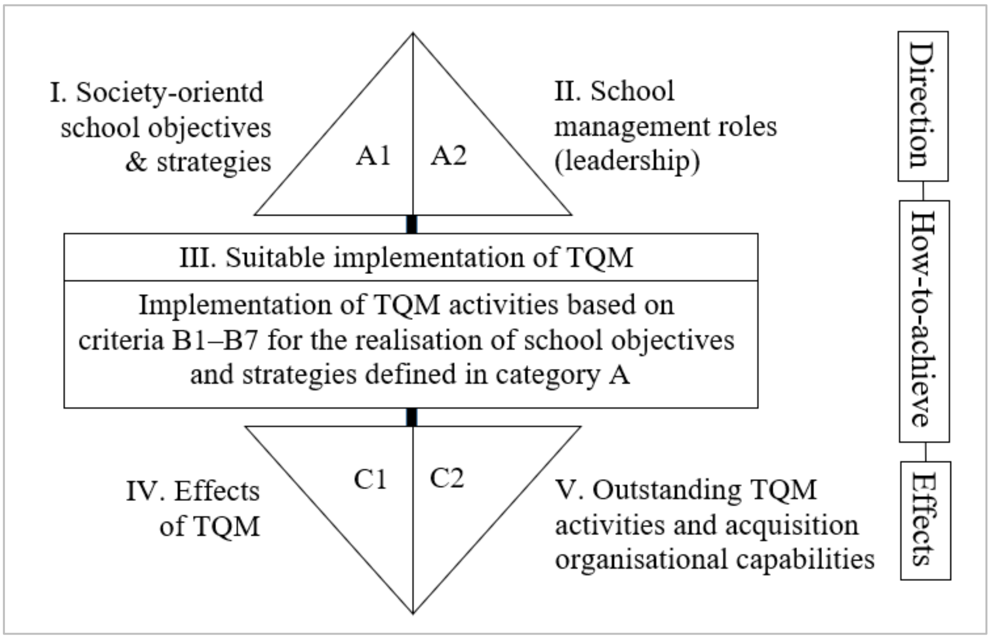


Figure 2. 1: Diagram of the TQM model based on Deming criteria for school education

Table 2. 1: Descriptions of each criterion of the TQM model based on Deming criteria to enhance organizational ability

Category A	
A1	Establishment of proactive society-oriented school objectives and strategies (Establishment of proactive customer-oriented business objectives and strategies) Formulation and establishment of society-oriented school objectives and strategies as school missions and visions; the school's future plans are clearly shared.
A2	Role of school top management and its fulfilment (Role of top management and its fulfilment) School top management, including the principal, are involved in formulating society-oriented school objectives and strategies, enhancement of school capabilities, and demonstration of enthusiasm for TQM implementation.
Category B	
B1	School-wide deployment of school objectives and strategies (Organizational deployment of business objectives and strategies) A systematic deployment of school objectives and strategies to achieve school excellence based on challenging targets. Deployment of objectives and strategies throughout the school organization based on cooperation among departments, faculties, and staff.
B2	Creation of new educational values based on understanding of societal needs (Creation of new values based on understanding of customer and social needs and innovation of technology business model) Development and management of new educational services and school processes are proactively and effectively performed, aiming to create new educational values based on societal needs. Innovative technologies are utilized for the creation of new values in education.
B3a	Management of quality in educational services and supporting work (Management and improvement of quality of products and services and/or work process) Implementation of subject-specific supporting work, educational management, and standardization of the school's routine work, especially in the core processes of education (teaching and learning). These activities are important for stabilizing the process of achieving the desired targets. This includes proper education and training for teachers and school staff in the provision of education to students.

B3b	<u>Continuous improvement in quality of all school services and processes</u> (Management and improvement of quality of products and services and/or work process) Implementation of continuous improvement in all school services and processes to systematically improve the quality of the school. Through continuous improvement activities, problematic areas in the provision of services, length of execution of routine tasks, and complaints from external parties may be reduced. This may increase the level of service satisfaction and happiness resulting from all school processes.
B4	<u>Establishment of school-wide quality assurance system</u> (Establishment and operation of cross-functional management systems such as quality, quantity, delivery, cost, safety, environment, and so forth across the supply chain) Establishment of a school-wide quality assurance system to improve overall education quality, especially in the provision of values in learning. For teaching staff, this covers the process of student admissions, placement in classrooms, and performance monitoring and graduation. Various departments must be involved to assure the highest quality of learning for students. The management system for operations related to safety, cost, environment, and school suppliers should also be involved.
B5	<u>Collection and analysis of information and utilization of accumulated knowledge</u> (Collection and analysis of information and accumulation and utilization of knowledge) Establishment of a system to systematically collect and analyse feedback from members of society. This also extends to the voices of parents, students, and school staff, and includes analysis of student performance and utilization of information technology (IT). Such information and accumulated knowledge are employed to create new values in educational services and operations to improve learning quality.
B6	<u>Development of school staff and proactive utilization of school capabilities</u> (Development and active utilization of human resource and organizational capability) Development of skills and relevant abilities in teachers and school staff in supporting departments in order to provide better education and embed a culture of quality at the school. Furthermore, school capabilities are used to realize school objectives and strategies in the TQM implementation.
B7	<u>Initiating and fulfilling the school's social responsibility</u> (Initiatives for social responsibility of the organization) Being active in performing school roles and obligations as part of society through social responsibility activities that contribute to the well-being of the immediate community. The school plays a role in advocating environmental sustainability and promoting ethical operating practices such as transparency and integrity.

Category C	
C1	<u>Realization of the effects of objectives and strategies through TQM implementation</u> (Effects obtained regarding business objectives and strategies through utilization and implementation of TQM) The school has realized the effects of its society-oriented school objectives and strategies through the implementation and utilisation of TQM suitable to the school context.
C2	<u>Exceptional TQM activities and acquisition of school capabilities</u> (Outstanding TQM activities and acquisition of organizational capabilities) Through exceptional TQM implementation, the school has had positive effects on its core processes of teaching and learning to realize school objectives and strategies and to acquire capabilities necessary for future growth.

Note: The original evaluation items from the general framework of the Deming Prize are written in parentheses.

2.3.2.2 Society-oriented school objectives and strategies in category A

For category A criteria, this chapter explores society-oriented school objectives and strategies. It is critical to define school objectives and strategies based on a society-oriented concept rather than using customer-oriented terminology. Potential stakeholders or school beneficiaries are listed based on guidance from relevant literature. Using the compiled list, the society-oriented stakeholders are classified into six categories and numbered according to importance: 1) internal; 2) parties who are frequently involved with school; 3) regulatory bodies; 4) surroundings; 5) immediate organizations after completion of education; and 6) other related parties. The descriptions of these society-oriented stakeholders indicate the value provided by the school to society and vice-versa, as shown by the solid and dotted lines in Figure 2.2. These values can be classified as tangible values, such as monetary contributions, and intangible values, such as knowledge. Process to determine Figure 2.2 can be referred to Appendix 1.

Education is primarily about teacher-student communication, especially within the daily teaching and learning process of transferring knowledge. As such, these are the parties with the most frequent interactions in the provision-of-learning process.

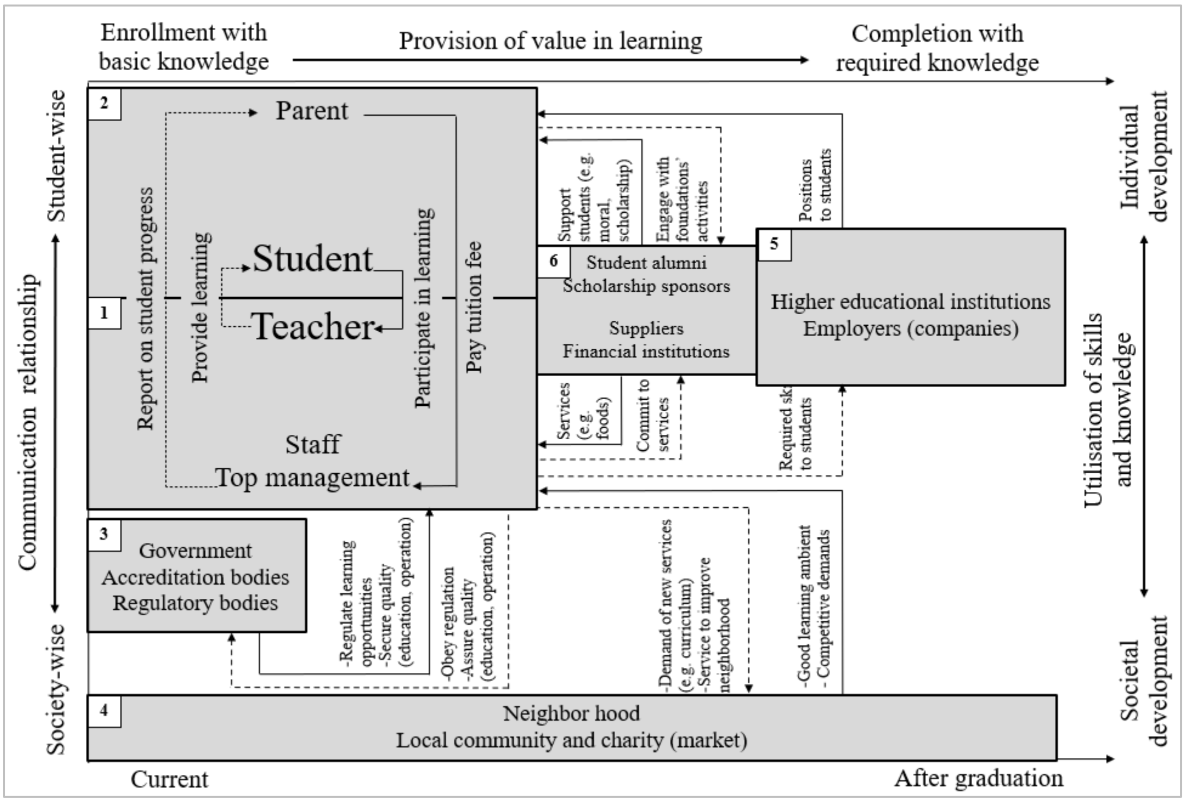


Figure 2. 2: Society-oriented parties/stakeholders to be considered in school objectives and strategies

2.3.2.3 Development of sub-criteria for categories A, B, and C

Based on the output from Step 1 (Table 2.1), sub-criteria in categories A, B, and C are developed. The method for establishing these sub-criteria is described below, taking sub-criteria for B1 as an illustrative example.

Step 2: Selection of literature related to each criterion

For B1, studies related to 'deployment of school objectives, strategies and policies' or 'implementation of school mission and vision' are identified. One source is the Guidelines for Policy Management published by the Japanese Society for Quality Control (JSQC), which provides practical guidelines for all types of organizations.

Step 3: Identification of factors for each criterion

Using JSQC (2017), factors such as a) formulation of policy by leaders, b) deployment of the policies to entire organization, c) prioritization of aims with appropriate means, d) execution of planning, e) review of implementation status, f) analysis of achieved and unachieved results, g) remedial action taken and h) term-end review are identified. These factors are then used to develop the B1 sub-criteria.

Step 4: Finalize sub-criteria and rephrase into educational terminology

First, from the JSQC list, factor a) 'formulation of policies by leaders', is used as the basis for sub-criterion B1(i). This factor is converted into an educational context as 'society-oriented school objectives and strategies are formulated and established by school principals, department heads, faculty heads, and coordinators based on school missions and visions.' Next, item b) 'deployment of policies to entire organization', suggests that schools should consider the deployment of objectives across all levels of the school organization with broad and open discussions to align horizontal and vertical levels. This is rephrased as sub-criterion (ii). The remaining factors c)–h) are converted to sub-criteria (iii) to (vi) in a similar manner. The full set of sub-criteria for B1 is shown in Table 2.2.

The content validity of all criteria and sub-criteria are checked. The identification, selection and framing of the criteria and sub-criteria are made through several stages. The engagement with an expert of TQM in education field is carried out to comment on the representativeness and suitability of the selected sub-criteria. Some weaknesses in the initial model are identified and suggestions are given to improve the statements for better understanding. Finally, the criteria and sub-criteria are finalized as a model that is useful to school management. Full versions of the developed criteria and sub-criteria, and a comparison with the general Deming Prize framework can be obtained at Appendix 2.

Table 2. 2: Descriptions of the B1 criterion and its sub-criteria

Criterion	Sub-criteria
B1: School-wide deployment of school objectives and strategies	(i) Society-oriented school objectives and strategies are formulated and established by the school principal, department heads, faculty heads, and coordinators based on school missions and visions.
	(ii) Society-oriented school objectives and strategies are deployed to the related department or faculty, vertically and horizontally, throughout the school.
	(iii) The main objective is divided into smaller objectives that are deployed with the appropriate methods and target values (numerical or measurable values, if possible).

	(iv) Review, analysis, and investigation of the current implementation status and results are carried out to clarify root causes of any unachieved results.
	(v) Appropriate remedial actions (corrections and correctives) are carried out based on the review and reflection outcomes.
	(vi) Periodic review is carried out (e.g., monthly, quarterly, half-year, yearly, at term end) led by department or faculty heads and coordinators.

2.4 Application of the model

2.4.1 Research design for data collection

A case study is conducted to evaluate the effectiveness of the proposed model. This qualitative research approach enables this study to explore the meanings that relevant groups or individuals ascribe to a social problem (Creswell, 2014). The subject of our case study is the Global Indian International School, Tokyo Campus, which is an international school in Japan that has embarked on a TQM implementation journey to achieve school excellence. In total, the school has 22 campuses worldwide, and is headquartered in Singapore. The Tokyo Campus was opened in 2006. It aims to be a leading international school that provides an international curriculum to students from diverse backgrounds. The school offers numerous educational programs and curricula, including the Central Board of Secondary Education and Cambridge International General Certificate of Secondary Education. In 2020, the school had nearly 820 students, from pre-school to senior high school. The school established its quality management system in 2018 upon receiving its ISO certification (9001:2015).

To clarify the current status of the school, a combination of data collection methods was used, including document inspections and direct observations. Firstly, a series of training sessions on our model was conducted on site to explain the substance of the TQM criteria, how to use the model and how to judge current practices in order to draft proper improvement plans. Following this comprehensive introduction, activities based on the model were implemented. The assessment was carried out for all model criteria in two phases (Phases 1 and 2) by a special team. A special team consisting of six people was formed, comprising two consultants for TQM implementation in the school and four members of school management personnel. The school management personnel have more than ten years of experience in managing education. In particular, they have full access to the case study school's systems and records in order to assess the current practice and present implementation evidence. They play a vital role in coordinating and communicating with the consultants and have been briefed on the application of the proposed model numerous times. As such, they have a thorough understanding of what are the requirements and how to apply the model based on criteria and sub-criteria. Besides, one of the personnel has authority in decision making for the direction of future improvements of the school. On the other hand, both consultants belong to an educational institution and have in-depth knowledge of TQM's principles and practices. Thus, it makes it possible for them to provide detailed guidance to school management personnel when they are deciding how to assign evaluation points based on sub-criteria. The evaluation scores for each sub-criterion in Table 2.1 ranged from 0 (no activity) to 10 (exceptional activity). TQM activities are considered exceptional (score of 10) when the activities are aligned with the school management philosophy, scale, and environment, and could be benchmarked against other schools.

To collect feedback on the model's effectiveness, two separate feedback collection methods was employed. Semi-structured interviews were conducted with the school principal and educational excellence managers using open-ended questions. The interviews were conducted in three independent sessions for each interviewee using the interview questions Iq1 to Iq4 (see Appendix 3). For the school's middle management and staff, a survey consisting of a combination of closed-ended and open-ended questions was administered. The open-ended and closed-ended survey questions are also included in the Appendix 3. Quantitative data from the closed-ended questions were used to determine the responses and provide a quick overall measure of the understanding of TQM implementation at the school. Ten responses were collected to the survey.

2.4.2 Findings

2.4.2.1 Implementation of the total quality management model based on Deming criteria

Several activities were conducted for each criterion. For the sake of conciseness, an in-depth assessment of the B1 criterion (implementation of school objectives and strategies) is presented in this chapter. Evaluation scores were assigned for each sub-criterion in Phases 1 and 2 shown in Figure 2.3. To determine whether the model's sub-criteria were fulfilled, Phase 1 assessment was performed with a focus on a recent major project related to the school's mission. The project relates to transforming the school into a multi-cultural and multi-curricular school. As such, it is a challenging project for the entire organization. This is one of the reasons for the needs of TQM implementation in the school, which is to ensure this project is successfully completed. Phase 2 is the stage at which a target improvement plan is decided upon, which mainly depends on the school principal's decision based on the school's resources.

For Phase 1, the evaluation score for sub-criterion B1(i) was judged as 7, based on the following factors: (a) society-oriented school objectives based on school mission and vision, (b) a process to determine objectives and strategies, and (c) top management's involvement in the formulation of objectives. The observed facts are as follows: (a) the school defined the school's objective as introducing new curricula with the aim of becoming a multi-curricular and multicultural school. This objective can be regarded as linked to the school's mission and vision 'to be a global role model in teaching and learning'; (b) a draft of the objective was decided in scheduled meetings, and the objective was then finalized in a higher-level meeting; and (c) it was confirmed that the top management leadership, including the school principal, were involved in the formulation of the objective. Although the process to determine objectives was generally established and shared among relevant individuals, some of the detailed procedures were not documented in the official work reports. For Phase 2, the target improvement level was set to 8, with a focus on the negative points in Phase 1.

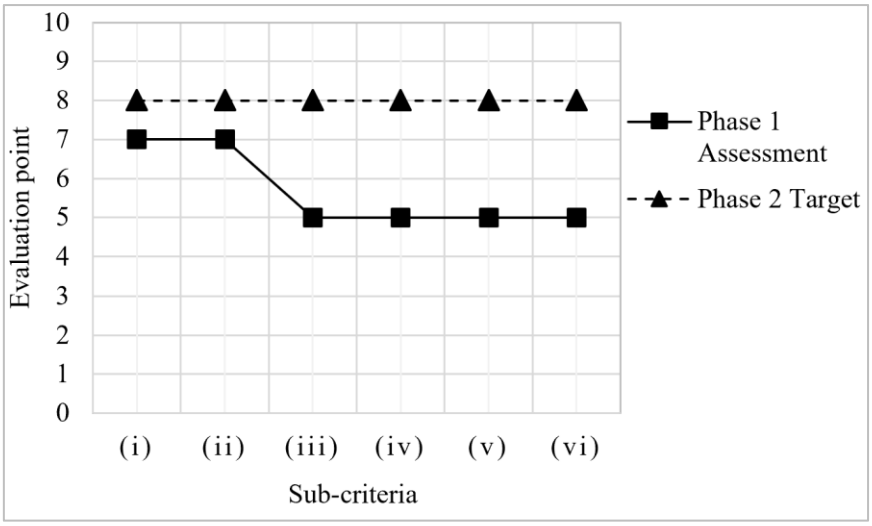


Figure 2. 3: Evaluation scores for the B1 sub-criteria

In a similar manner, the evaluation score was decided as 5 for sub-criterion B1(iii) in Phase 1, based on the following facts: (a) the objective of introducing the new curricula with the aim of becoming a multi-curricular and multicultural school was deployed as a set of smaller objectives, such as the implementation of another international curriculum; (b) the time target is confirmed to be two years; and (c) the means to achieve the target were not clearly determined. For Phase 2, the target improvement level was set to 8. Hence, the school needs to carry out major improvements, primarily to clarify the means by which they will achieve their target. Moreover, the person in charge needs to be thoroughly briefed and trained on the proposed means to ensure that that process is performed correctly.

2.4.2.2 Feedback on the model's effectiveness

(i) Enhancing organizational ability

The school principal and the school's educational excellence managers X and Y agreed that the TQM model based on Deming criteria promoted organizational sustainability and agility at the school. They stated that they experienced this benefit to a significant extent. Based on Iq1, educational excellence manager X stated that the school had also adopted the ISO 9001 and Baldrige model, and that utilization of the TQM model based on Deming criteria provided additional benefits and processes that were well defined. The manager explained that 'this model, based on A, B and C category, is definitely suitable for the school's future stability'.

(ii) Society-oriented school objectives and strategies

Iq2 and Iq3 (see Appendix 3) were asked during the interview sessions to identify the abovementioned benefits. The responses of the school principal and educational excellence managers agreed that the society-oriented objectives and strategies suggested effective ways to recognize appropriate stakeholders. Educational excellence manager X noted that seeing the size of the boxes and their location on the horizontal and vertical axes encouraged the school to give more consideration in those areas.

(iii) Implementation of TQM activities and the model's overall usage

Responding to Iq4, the school principal and educational excellence managers agreed that, particularly in the implementation of school objectives and strategies under criterion B1, the sub-criteria served as effective guidelines; for example, sub-criterion B1(iv), which specified 'review, analysis, and investigation of the current implementation status and results are carried out to clarify root causes of any unachieved results', was deemed effective. Accordingly, respondents agreed that the achievement status needs to be checked to determine whether it is still on track or diverges from the target plan. This sub-criterion, in the view of the school principal and educational excellence managers, provided hints to school management regarding the specific provision of means to systematically execute required processes to introduce new curricula. The school principal speculated that if the school closely followed the sub-criteria, it would have a robust implementation system for achieving its targets.

In addition to the interviews, completed surveys on the model's effectiveness were received from ten respondents from the school's middle management and staff. The questionnaire consisted of two questions, Sq1 and Sq2. Five respondents answered that they 'agree' or 'strongly agree' with the statements, while the remaining answered that they 'neither agree nor disagree'. No one answered 'disagree' or 'strongly disagree'. Particularly with respect to teaching and learning work in the classroom, one staff member agreed that the TQM model based on Deming criteria raised the visibility of school strengths and weaknesses, as the model identified and established proper processes to accomplish the required improvement and enhanced the monitoring of lower performers in the assessments.

(iv) Comparisons with other models

Educational excellence manager Y explained her view on the difference between the Baldrige framework and the TQM model based on Deming criteria. The manager stated that, based on the school's implementation of the Baldrige framework over a period of nearly twelve years, the Baldrige model provides a framework of criteria and overall assessment. The questions are straight-forward and intended to be used as guidelines. The detailed explanation and example of Baldrige Excellence Framework (Education) are described in Appendix 2. The Baldrige framework questions typically begin with 'how' and 'what' to assist applicants in explaining their processes in relation to the criterion. The prescriptive guideline helps the Baldrige examiners to assess the applicant when they seek for an external evaluation. On the contrary, due to the manager, the proposed TQM model emphasizes clarity of process in achieving results and offers a systematic way to implement appropriate activities. Adopting the proposed model helped management drill down further into the school process, and it contained more requirements in terms of what-to-achieve and how-to-achieve. In applying the proposed model, school managers were better able to understand how deep they needed to go and how precise they needed to be; this differed significantly from their experiences of applying the Baldrige model. To quote educational excellence manager Y, 'Deming is deeper and more granular. If Baldrige is a molecule, Deming is a nano-sized molecule.' The comparison between Baldrige

Excellence Framework (Education) and the TQM model based on Deming criteria is highlighted at Appendix 2. Based on the comparison, the proposed model which is inspired by the general Deming framework's philosophy includes the extensions of sub-criteria for each criterion to guide schools on how to conduct TQM activities in a systematic way. The general Deming framework places emphasis on to-be-status. The organizations are expected to understand their current condition, set their own themes and goals, and transform the organization. This explains why the Baldrige framework is more prescriptive than the Deming framework.

2.5 Discussion

2.5.1 Insights from the findings and advantages of the model

The original comments are almost duplication with 2.4.2, without insights. As for the comparison between models, the ISO standard and Baldrige Excellence Framework (Education) encourage organizations to utilize methods such as balanced scorecards and plan-do-check-act (PDCA) to achieve their objectives. The TQM model based on Deming criteria focuses on both what-to-achieve and how-to-achieve through the use of a systematic management approach incorporating appropriate statistical analysis of its implementation. In this model, the implementation of TQM activities in all areas is based on a systematic management cycle. Having such sub-criteria as a guide to improving weaknesses in school practices offers different views to school management regarding how to systematically manage the TQM criteria with suitable process analysis. Thus, a school will be equipped with the ability to face challenges and to increase its quality to a higher level in the future. Additionally, the proposed TQM model recognizes clarity in the process of achieving results. It assists schools in gaining greater awareness of their processes through the adoption of sub-criteria in various phases and improvement plans. Interviewees agreed that this occurred during the model's implementation.

The TQM model based on Deming criteria underlines the importance of school objectives and strategies associated with society-oriented goals. Through the appropriate recognition of how society relates to the school, the model promotes a balanced recognition of the various stakeholders in a way that engenders their support in realizing school objectives and strategies. Moreover, the school is able to recognize the role of not only school-related external parties, but also internal personnel who manage and support the organization and provide the necessary foundation for future school growth.

2.5.2 Concluding remarks

Based on the general Deming Prize criteria, this chapter constructed a TQM model for achieving school excellence. The model is developed in accordance with the Deming Prize's A, B and C categories as an overall framework. To ensure systematic practices in the implementation of the TQM model, sub-criteria for each category are formulated based on multiple factors extracted from the relevant literature. These factors are assessed and rephrased as sub-criteria that suit the school setting. The criteria and sub-criteria for what-to-achieve and how-to-achieve function as guidance for identifying and improving weaknesses in school practice. Through implementation of the proposed model, schools will be equipped with the ability to overcome hurdles and increase the quality of learning and the effectiveness of their operations.

The model replaces the customer-oriented terminology of the Deming Prize criteria with society-oriented terminology and principles. The society-oriented school objectives and strategies are unified in the model's category A. Stakeholders related to the school are listed and classified into six categories to represent the various society-oriented parties and stakeholders in the formulation of the school's objectives and strategies. By implementing society-oriented school objectives and strategies and linking them with results, schools can appropriately balance the interests, priorities, and contributions of the important parties and stakeholders as a means for gaining their support in the realization of their objectives.

2.5.3 Implication of study

On a practical level, the model's adoption enhances organizational ability in education management, not only for the present but also for future advancement towards school excellence. Furthermore, because the formulation of appropriate school objectives and strategies is one of the important inputs of the model, it encourages school management to create a balanced recognition of all stakeholder interests in order to gain support to realize the school's objectives. Additionally, school management is able to connect the results and impacts of TQM implementation to the objectives and strategies that are formulated in the initial stage.

Also, teachers, school staffs and management may use the proposed model as a guideline to perform each TQM activity systematically. Moreover, the model may be used as a self-assessment tool to assess the current practice. By doing so, it allows schools to carry out improvements for each sub-criterion based on evaluation score targeted in the next stage.

2.5.4 Limitations

This chapter has some limitations. Firstly, it uses data from one school as a case study to validate the model's effectiveness. In the future, the model should be implemented in other schools using the same study methods. Secondly, the implementation of a longitudinal research design to obtain evidence of the model's effectiveness over time would strengthen evidence for the model's effectiveness. Thirdly, the weightings for all criteria in the model's B category are equal, with the same scores ranging from 0 to 10. This is done to familiarize the special team with the scoring ranges for each activity. Furthermore, future studies should seek feedback from educational institutions that are presently adopting TQM practices based on Deming Prize criteria or MBNQA award winners to enable more comparison between models. Additionally, future research should undertake a detailed comparison between EFQM for education and the TQM model based on Deming criteria, as both models have some general commonalities. Despite these limitations, it is hoped that the proposed model can help advance the field of knowledge in the important area of TQM in schools and other educational institutions.

CHAPTER 3: Detection of unusual scores of individual students for immediate remedy according to daily assessment data

3.1 Introduction

The educational institutions cope with the competitive world by providing high-quality education and continuously improving their processes (Venkatraman, 2007). Total quality management (TQM) has been extensively adopted in industries and service sectors (Bayazit, 2003; Kumar *et al*., 2011; Yapa, 2012; Verma *et al*., 2022). Lately, it is increasingly being adopted in educational institutions (Sahney *et al*., 2004; Töremen *et al*., 2009; Nawelwa *et al*., 2015), including teaching and learning (Crawford and Shutler, 1999; Mehra and Rhee, 2004). One of the important TQM components is daily management. It is widely adopted in numerous sectors to stabilize results in day-to-day jobs and routine activities. Japanese Society of Quality Control (JSQC) opines that the daily management practice is essential in organizations to ensure that the process is carried out according to its standards efficiently to achieve the objectives (JSQC, 2014). It is a system that documents the processes related to a particular job and the plans for managing and improving operations (Cassidy, 1996). Daily management has been practiced not only in industries (Zarbo *et al*., 2015; Kennedy, 2019) but also in service sectors (Donnelly, 2014; Whitley *et al*., 2015; Hopkins *et al*., 2017; Wirtz *et al*., 2021).

Daily management establishes the processes and systems that can achieve a certain level of stability. This practice is used for maintaining good results by following standard operating procedures. Hence, whenever a completed job deviates from the target, the management should assess the process standardization status, such as inspecting whether standard operating procedures are followed (e.g., Ohno and Bodek, 1988; JSQC, 2014). There are various activities included in daily management such as clarification of daily process flow, process standardization, detection of abnormality, immediate remedies, and actions for the root cause. In daily activities, the results of a process may fluctuate due to a variety of causes. The unfavorable phenomena shall be removed by immediate remedies. The utilization of statistical methods to detect the existence of deviation of operation from the process standard on the assignable causes is broadly applied in industries, and is gradually being adopted in education systems.

Information technology (IT) systems adopted in schools collect various types of data, such as the students' demographics, financial operations, student test scores and satisfaction levels (Ikemoto and Marsh, 2007). In the past, daily assessment data were often stored in teachers' computers individually. They tended to report the final scores before the end of the semester for grading purposes. Recently, because of the progression of IT systems, student performance data, based on daily assessments, are available through centralized monitoring systems in some schools. The teacher inputs these data into the system after the students' assessments. They can be viewed not only by the teacher but also by the school's management. Student scores in various subjects, assessments, and student cohorts are shared. Data utilization can be of benefit in detecting abnormality considering student ability, subject difficulty, and related factors which influence student scores. This would lead to an immediate remedy taken in the same academic year.

In this chapter, the terminology of unusual scores is used to represent the occurrence of abnormality in student scores. This chapter focuses on the detection of unusual scores for individual students by applying a linear ANOVA model. Daily assessment data is used to evaluate the students' understanding of what they have been taught. After students sit for their daily assessment of a particular subject, the scores of individual students are determined and stored in a centralized monitoring system. An unusual score is a result of deviation from the normal stable situation of an individual student when the teacher evaluates the student based on daily assessments. The unusual score is detected by the model after the daily assessment of different subjects for each student. The detection which is below the lower limit provides hints for an immediate remedy to improve the quality of teaching and learning by taking appropriate actions, considering assessment results and processes to determine such results. Whereas the detection that is beyond the upper limit indicates a good practice in learning and teaching based on a truly outstanding score. A dataset from an international school in Japan is utilized as a case study for model validation.

3.2 Literature Reviews

3.2.1 Daily management in education

Numerous educational systems are in place to monitor student progression for better decision-making. It also offers specific guidance to teachers on how to execute the process. A standardized curriculum and syllabus promote a sense of structure for a school to provide common instruments to deliver instructions and assess students' understanding (Murgatroyd, 1992; Scheerens, 2016). Following such standards in the teaching and learning process is a fundamental practice of daily management. Formative and summative assessments are two tools that teachers commonly employ to assess the students' learning of new material and knowledge of standards. Formative assessments are activities undertaken by teachers and their students in evaluating themselves, which provide the information which is utilized as feedback to modify teaching and learning activities. It should be used during instructions to support student learning material throughout the learning process. On the other hand, summative assessments are cumulative assessments that intend to capture what a student has learned or the quality of the learning, and judge performance against some standards. It is used at the end of a unit, chapter, or semester to evaluate how much learning that students have gained (Miller *et al.*, 2009; Dixson and Worrell, 2016; Connors, 2021).

Detection of abnormality is one of the essential elements in daily management. The fundamental concept evolved from the statistical control chart by Walter A. Shewhart. Shewhart (1931) identifies two broad causes of variations; chance (known as common) causes, and assignable causes. The unknown cause of variability in the quality of a product that does not belong to a constant system is known as an assignable cause. An abnormality is an event in which a result of the process deviates from the normal state because of variations in assignable causes. JSQC (2014) highlights that abnormality should be clearly distinguished from non-conformance. The terminology of non-conformance is used when a product, service, or process does not meet the specified requirements. In education, the terminology of non-conformance could be signified as failure or at-risk (Kaufman and Bradbury, 1992). For example, the at-risk score may be represented by a threshold value, which is a cut-off score that a student needs to obtain in an assessment or course. The school management has to realize that in controlled conditions, obtaining an at-risk score is considered normal when standards are followed. In other words, if the threshold value is 40, it is normal if a student achieves a score of below 40 in other assessments when teachers and students follow the standards.

Abnormality represents data where the process is not in a controlled state. A controlled state is when a process is stable in achieving the desired target and is called a normal condition. Four scenarios may occur in a process with individual judgments of output. The categories are a) when both non-conformance and abnormality do not exist, b) when non-conformance does not exist, but abnormality exists, c) when non-conformance exists, but abnormality does not exist, and d) when both non-conformance and abnormality exist. The data obtained in the daily process are plotted in a control chart which includes the control limit to identify the existence of abnormality. If one or more points exceed the control limits or have a specific trend, assignable causes shall be considered to exist and immediate actions must be taken to prevent the recurrence (e.g., JSQC, 2014). Although literally, abnormality is an opposite term of normal, it is still difficult to relate 'abnormality' in the context of teaching and learning because it is likely to invite a negative perception. Thus, using 'unstable' or 'unusual' would be more appropriate in the context of education. For instance, a stable condition is when a student's performance, represented by the assessment scores, is within the normal range. Data that fluctuates from the normal region is considered an unusual score. A daily management model for teaching and learning for school education can be referred at Appendix 4.

3.2.2 Statistical data analysis on educational scores

Lately, the usage of statistical process control tools has been getting attention in the field of education. Student scores are utilized to determine the existence of unusual process variations in courses offered by educational institutions. Savic (2006) utilizes control charts to analyze the existence of variations in the existing grading process to determine if it is compatible with the credit transfer system requirements. Five types of passing grades during nine examination periods in the same academic year are used to plot the p-chart. The plots beyond the control limit are treated as unusual variations in the nine periods that exist for each subject.

Hanna *et al*. (2012) highlights the existence of special causes of variations in passing rates of the courses in future semesters. A p-chart is used to determine the average proportion of selected courses such as mathematics, business statistics, and management. The proportion of students achieving a grade of C or better is calculated for each of the five courses during spring and fall semesters within nine years. The variation in passing rates provides hints to academic administrators who seek to improve retention, progression, and graduation rates.

Beshah (2012) applies a $\bar{X} - R$ chart with control limits to devise a new method of grading system to assign grades more objectively. The analysis data is obtained from an Engineering Mechanics (Dynamics) course at an institute of technology. Dataset is obtained from twenty groups in a course with a size of approximately forty students in each group. Five students' scores from each group are randomly selected to obtain the average and range for each group. Each upper and lower control limit determined from the $\bar{X}$ chart is utilized as a grade interval for the letter grades ranging from A to F. It is an appropriate approach to evaluate the group average and variation within each group.

Huang and Fang (2013) compare four types of predictive mathematical models to predict student performances in a highly enrolled core course. Data are collected from a total of 323 students from different majors who have enrolled for a course. The individual student data contains the cumulative grade point average for four semesters, grades earned in four pre-requisite courses, and scores on three dynamic mid-term exams. The models can be utilized to predict a student's future performance.

In some studies, the research related to factors that may influence student scores is conducted to provide empirical evidence on the suitability of teaching and learning modes in courses. Alagiah *et al*. (2001) studies the relationship between first-year student performances and their attendance at tutorials. Syukur (2021) examines whether the student's achievement differs across their gender, study major, and origins. Moreover, Otavová and Sýkorová (2016) explore the differences among test scores of mathematics for economists courses obtained by students under five different faculties. The dataset contains the scores for 2,256 students in a course, and the main factors of faculty (five levels) and semester (two levels) with their interaction are analyzed using a linear model.

3.3 A model for detection of unusual score for daily assessment

3.3.1 Daily assessments data

The fundamental idea behind the proposed model is to detect unusual scores of an individual student just after the assessment is carried out in the same academic year. It is to prevent problems before going to the next level within the same academic year, once the score which deviates from a normal state is detected. The unusual score can be interpreted based on two conditions: i) unusual (low) score to represent a truly critical score, and ii) unusual (high) score to describe a truly outstanding score obtained by students. The dataset includes all individual students' scores based on several subjects and assessments. The students belong to the same classroom, learning the same subjects, and sitting for the same assessments in one academic year.

The data are collected and managed through the schools' centralized monitoring systems. The advancement of IT systems in schools enables the management to integrate students' scores for all types of assessments and allows immediate remedy to be taken after an assessment, considering student ability, subject difficulty, and so forth. The daily assessment activity is performed based on the academic subject, such as science, English, and mathematics. The assessments are mainly based on formative assessments carried out in the classroom. The assessment grading uses a numerical grading system such as from 0 to 100. Monthly assessments, such as assignments of individual activities, group activities, notebook maintenance, and pen and paper tests like multiple-choice question (MCQ) tests and quizzes, are carried out regularly.

Subject: Science ($j = 3$)										
Student (i)	Assessment (t)									
	1	2	3	4	5	6	7	8	9	10
A										
B										

Subject: English ($j = 2$)										
Student (i)	Assessment (t)									
	1	2	3	4	5	6	7	8	9	10
A										
B										

Subject: Mathematics ($j = 1$)										
Student (i)	Assessment (t)									
	1	2	3	4	5	6	7	8	9	10
A										
B										
•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•
K										

Figure 3. 1: Outline of the dataset for daily assessments

Figure 3.1 shows the outline of the dataset for daily assessments compatible with the model developed in this chapter. The figure describes the dataset of assessment scores based on three factors: Student i's score in subject j on the t-th assessment. The subject j are the scholastic subjects such as mathematics, English, and science. The process for detection of the unusual score is carried out after an assessment is completed. For example, Assessment 4 is planned to be carried out in April. When an unusual (low) score of a student is detected after

Assessment 4 in April, the teacher must take an immediate action to remedy an arising problem, such as supplemental lectures, before proceeding to Assessment 5.

3.3.2 A model for detection of unusual scores of an individual student

3.3.2.1 A model for data analysis

To take action for the student whose score is unusual, a linear model is considered such that student i obtains score Y_{ijt} for subject j on the t-th assessment, where $i = 1, \dots, n$, $j = 1, \dots, S$ and $t = 1, \dots, T$. It is assumed that Y_{ijt} can be expressed by

$$Y_{ijt} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + (\beta\gamma)_{jt} + \varepsilon_{ijt}, \qquad (3.1)$$

where μ is a general mean of assessment score, α_i is an ability of student i, β_j is a difficulty of subject j, $(\alpha\beta)_{ij}$ is an ability of student i in subject j, and $(\beta\gamma)_{jt}$ is a difficulty of t-th assessment for subject j. Additionally, error term ε_{ijt} is included in the assessment score Y_{ijt}. It can be regarded as an application of three-factor models that is utilized ANOVA for statistical analysis. Detailed explanations of ANOVA are provided in standard textbooks on design of experiments, such as Montgomery (2012).

The purpose is to detect unusual score Y_{ijt} just after the T-th assessment because immediate remedy for the student whose score is unusual is crucial; this is like the principle of daily management in general. Let $\hat{Y}_{ijt}$ be an estimate, obtained by least square estimates of $\hat{\mu}, \hat{\alpha}_i, \hat{\beta}_j, \left(\widehat{\alpha\beta}\right)_{ij}, \left(\widehat{\beta\gamma}\right)_{jt}$ such that

$$\hat{Y}_{ijt} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \left(\widehat{\alpha\beta}\right)_{ij} + \left(\widehat{\beta\gamma}\right)_{jt}. \qquad (3.2)$$

Based on the residual $Y_{ijt} - \hat{Y}_{ijt}$, unusual (low) score identified when the residual is below the pre-specified value based on $\hat{\sigma}$ as

$$Y_{ijt} - \hat{Y}_{ijt} < -k\hat{\sigma}, \qquad (3.3a)$$

and unusual (high) score identified when the residual is beyond the pre-specified value based on $\hat{\sigma}$ as

$$Y_{ijt} - \hat{Y}_{ijt} > k\hat{\sigma}, \qquad (3.3b)$$

where k is a positive integer determined by management aspects considering type I and II errors. In addition, the estimate of error variance: $\hat{\sigma}$ is obtained by the root mean square of residuals $Y_{ijt} - \hat{Y}_{ijt}$. An unusual (low) score can be detected based on the Lower Control Limit

(LCL) and an unusual (high) score can be detected based on the Upper Control Limit (UCL), such that

$$LCL = \hat{Y}_{ijt} - k\hat{\sigma}. \quad (3.4a)$$

$$UCL = \hat{Y}_{ijt} + k\hat{\sigma}. \quad (3.4b)$$

3.3.2.2 List of symbols and estimates

Y_{ijt}: Score of student i for subject j on the t-th assessment;

$\hat{Y}_{ijt}$: Predicted score of student i for subject j on the t-th assessment by equation (2);

$\mu, \hat{\mu}$: General mean and its estimate;

$$\hat{\mu} = \sum_{t=1}^{T} \sum_{j=1}^{S} \sum_{i=1}^{n} Y_{ijt}/nTS \quad (3.5)$$

$\alpha, \hat{\alpha}_i$: Ability of student i and its estimate;

$$\hat{\alpha}_i = \sum_{j=1}^{S} \sum_{t=1}^{T} Y_{ijt}/ST - \hat{\mu} \quad (3.6)$$

$\beta, \hat{\beta}_j$: Difficulty of subject j and its estimate;

$$\hat{\beta}_j = \sum_{i=1}^{n} \sum_{t=1}^{T} Y_{ijt}/nT - \hat{\mu} \quad (3.7)$$

$(\alpha\beta)_{ij}, (\widehat{\alpha\beta})_{ij}$: An ability of student i in subject j and its estimate;

$$(\widehat{\alpha\beta})_{ij} = \sum_{t=1}^{T} Y_{ijt} / T - \sum_{j=1}^{S} \sum_{t=1}^{T} Y_{ijt}/ST - \sum_{i=1}^{n} \sum_{t=1}^{T} Y_{ijt}/nT + \hat{\mu} \quad (3.8)$$

$(\beta\gamma)_{jt}, (\widehat{\beta\gamma})_{jt}$: A difficulty of t-th assessment for subject j and its estimate;

$$(\widehat{\beta\gamma})_{jt} = \sum_{i=1}^{n} Y_{ijt} / n - \sum_{i=1}^{n} \sum_{t=1}^{T} Y_{ijt} / nT - \sum_{i=1}^{n} \sum_{j=1}^{S} Y_{ijt}/nS + \hat{\mu} \quad (3.9)$$

$\sigma, \hat{\sigma}$: Standard deviation of error and its estimate;

$$\hat{\sigma} = \sqrt{\left(\sum_{t=1}^{T} \sum_{j=1}^{S} \sum_{i=1}^{n} (Y_{ijt} - \hat{Y}_{ijt}) \right)^2 / (nST - nS - ST + S + T - 1)} \quad (3.10)$$

$\hat{Y}_{ijt} - k\hat{\sigma}$: Lower control limit (LCL)

$\hat{Y}_{ijt} + k\hat{\sigma}$: Upper control limit (UCL)

3.3.2.3 Detailed description of the model

When an unusual score is detected, a student's scores in the previous daily assessments are investigated to understand why their performance has suddenly deviated from the stable scores. As a description of each factor effect, α_i is interpreted as an ability of student i. Each student will have different abilities when compared to others, and it is a natural attribute of human beings. The term β_j is defined as the difficulty of subject j. If two academic subjects are chosen, for example, mathematics and English, the difficulty level for both subjects can be compared. For the interaction effect of $(\alpha\beta)_{ij}$, it can be described as the ability of student i in subject j. Based on the example of the two subjects, student i has a higher ability in mathematics, as compared to English. The interaction effect of $(\beta\gamma)_{jt}$ classifies a difficulty of t-th assessment for subject j. The t-th assessment is carried out independently for each subject. In this model, $t-$th assessment represents a repetition of the number of times for daily assessments to assess student understanding. The assumption is that the error variance of each assessment is independent of each other and not time-dependent.

The decision to control limit is an important process in the daily management practice to systematically manage the results that deviate from a normal trend in processes or systems. As example, for unusual (low) score, LCL is indicated to ensure that poor conditions that deviate from the normal state are immediately remedied. The selection of LCL through the selection of k by school management is based on the balance between type I and II errors. A type I error is when the proposed model detects an individual student who has a normal score. Students who work hard to achieve such results might get frustrated when they are detected as deviating from a normal trend, although they have done their best. Furthermore, in handling the matters, the school will be overloaded with problems that are not critical and will increase teachers' fatigue. On the other hand, a type II error is when the proposed model does not detect the unusual score of an individual student. Misidentifying a student who is obtaining an unusual score as normal prevents teachers from taking further actions to assist the student to get back to normal.

To determine k, schools need to analyze the different boundaries of LCL or UCL such as $k = 1, 2, 3$ for the dataset. In general, for example in lower boundaries, more detections of unusual (low) scores can be determined by $LCL = \hat{Y}_{ijt} - \hat{\sigma}$, and less detection for $LCL = \hat{Y}_{ijt} - 3\hat{\sigma}$. The normal distribution suggests the range between approximately 15.9%, 2.3% and 0.1% for unusual data lies below or beyond the $\hat{\sigma}, 2\hat{\sigma}$ and $3\hat{\sigma}$ range, respectively.

3.4 Case study

3.4.1 Outline of case

Data is obtained from the Global Indian International School, Tokyo Campus. This is an international school in Japan that is on its TQM implementation journey to achieve excellence.

The school has 22 campuses around the world, with headquarters in Singapore. The Tokyo Campus was established in 2006 and aims to be a leading school that provides an international curriculum to students from diverse backgrounds. The school offers several educational programs and curricula, including the Cambridge International General Certificate of Secondary Education and the Central Board of Secondary Education. In 2020, the school had almost 820 students, from pre-school to senior high school. The school established its quality management system in 2018 receiving its ISO certification (9001:2015).

Figure 3.1 demonstrates the selected dataset of 11 students, 3 subjects, namely, mathematics, English, and science, and 10 daily assessments in a semester. The assessments are formulated during the annual curriculum plan according to syllabus and methods using existing curricular standards. The analysis is simulating a condition where Assessment 10 for all subjects has just been completed. Then, teachers are managing the data to evaluate student understanding and scoring records.

3.4.2 Analysis results based on the model

The results of the ANOVA table based on the model in equation (3.1) are given in Table 3.1.

Table 3. 1: ANOVA table for dataset

Source	DF	Sum of Squares	Mean Square	F-ratio	Prob > F
Student: α_i	10	15947.86	1594.79	13.33	<.0001
Subject: β_j	2	245.99	122.99	1.03	0.36
Student×Subject: $(\alpha\beta)_{ij}$	20	4834.55	241.73	2.02	0.01
Assessment×Subject: $(\beta\gamma)_{jt}$	18	11100.56	616.70	5.16	<.0001
Error	279	33370.64	119.61		
Corrected Total	329	65499.59			

Notes: $R^2 = 0.49$, Adjusted $R^2 = 0.40$

Overall, Student, Student × Subject, and Assessment × Subject have statistically significant effects on student scores. The interaction of Student×Subject indicates that the relationship between student ability and scores depend on the difficulty level of a subject. Also, the interaction of Assessment×Subject indicates that the relationship between a difficulty of t-th assessment and scores depends on the subject. The $R^2 = 0.49$ suggests that nearly half of the variations of student scores can be explained by their general ability, their specific ability in a subject, and a difficulty of t-th assessment for a specific subject. Based on the analysis, there is no evidence showing that the error variance is not constant. Detail of analysis can be referred at Appendix 5.

Taking an example of student H in Assessment 10 for mathematics, $\hat{Y}_{H,1,10}$ is calculated. The parameter estimates for main effects of student H, subject of mathematics, a special ability of student H in mathematics, and difficulty level of Assessment 10 in mathematics are obtained using the ordinarily least squares method. The estimate of general mean is 86.74. $\hat{Y}_{H,1,10}$ is calculated as follows:

$$\hat{Y}_{H,1,10} = \hat{\mu} + \hat{\alpha}_H + \hat{\beta}_1 + \left(\widehat{\alpha\beta}\right)_{H,1} + \left(\widehat{\beta\gamma}\right)_{1,10}$$

$$= 86.74 + 1.69 - 1.15 - 4.98 + 6.31 = 88.61.$$

From the ANOVA analysis, $\hat{\sigma} = 10.94$, which is the estimate of error variance for this dataset. After confirmation with school management about the $k\hat{\sigma}$ boundary, they have chosen $k = 2$ considering the balanced of type I and type II errors to indicate LCL, because $k = 2$ provides an intermediate decision, neither too less nor too many detections of unusual score in this analysis. Therefore, it is calculated as follows:

$$LCL = \hat{Y}_{H,1,10} - 2\hat{\sigma} = 88.61 - (2 \times 10.94) = 66.73.$$

The actual score $Y_{H,1,10}$ is equal to 40. It shows that $Y_{H,1,10}$ is below LCL. It is judged as an unusual (low) score for student H. The same calculation is performed for $\hat{Y}_{i,j,10}$ to check for occurrence of unusual (low) and unusual (high) score for each assessment. The $Y_{i,j,10}$, $\hat{Y}_{i,j,10}$, LCL and UCL values for each student are plotted in Figure 3.2 and 3.3 for the two subjects, mathematics, and English, respectively.

Both graphs show the detection of unusual scores after Assessment 10 is completed, and before the students sit for Assessment 11. For student J, $Y_{J,1,10}$ is 40 as compared to its LCL, which is 54.5. There is no detection of unusual (low) scores for English in Figure 3.3. Also, there is no detection of unusual (high) score for all students in both subjects at Assessment 10. As such, the description in this chapter is focusing on the investigation of occurrence of unusual (low) score, while further explanation of detection of unusual (high) score is described in Appendix 8.

To further investigate the reasons behind the unusual (low) scores detected in Assessment 10 for student H and J, this chapter includes the analysis of both students' previous scores from Assessment 3 to Assessment 9. The LCL and UCL for each assessment is calculated based on the scores until t-th assessment, as shown in Figure 3.4. Each LCL and UCL is calculated based on t-th assessment separately. For example, for LCL in Assessment

5, the student scores during $t = 1, 2, 3, 4, 5$ are used for the calculation. Thus, different t-th assessment analysis produce different $\hat{\sigma}$ for calculation of LCL.

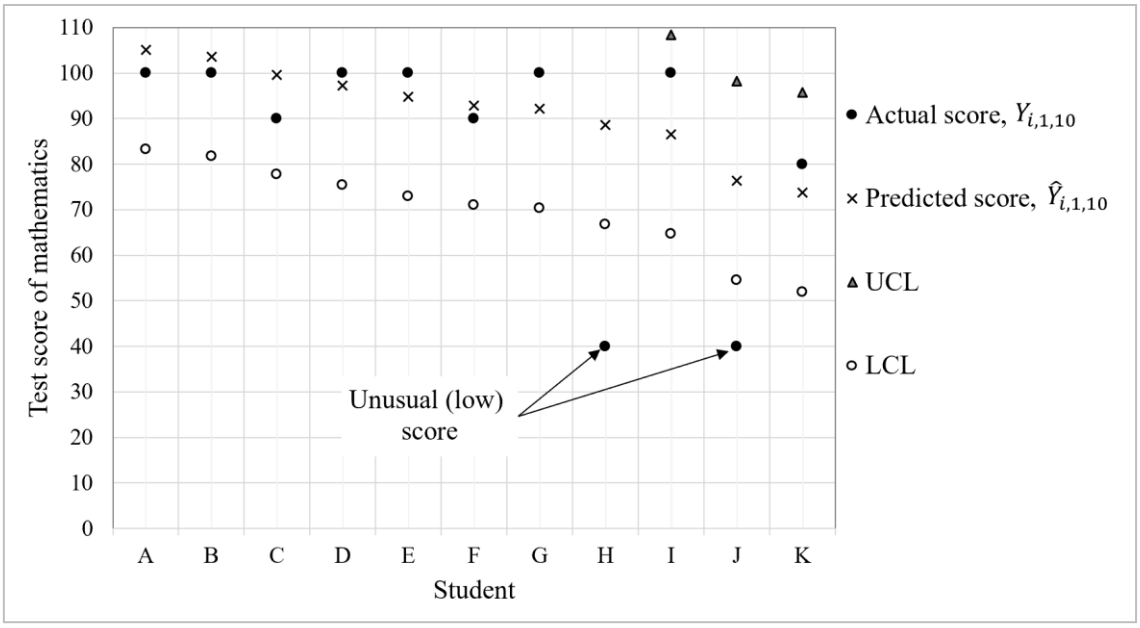


Figure 3. 2: Actual assessment scores against predicted scores, $\hat{\boldsymbol{Y}}_{\boldsymbol{i},\mathbf{1},\mathbf{10}}$, with LCL and UCL level of each student at Assessment 10 in mathematics

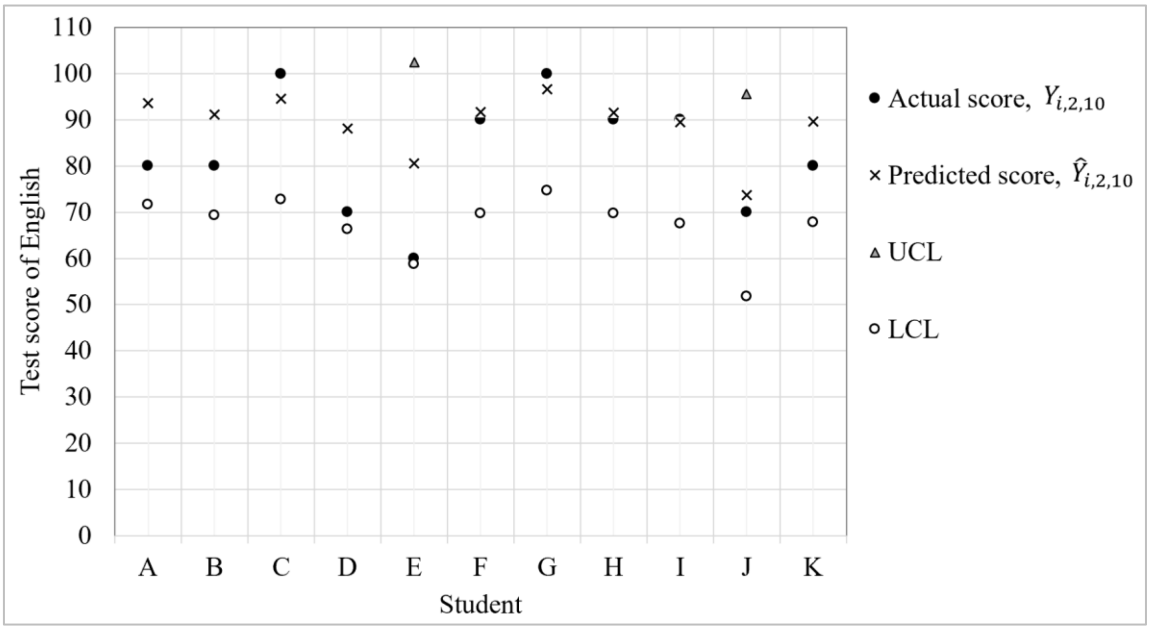


Figure 3. 3: Actual assessment scores against predicted scores, $\hat{\boldsymbol{Y}}_{\boldsymbol{i},\mathbf{2},\mathbf{10}}$, with LCL and UCL level of each student at Assessment 10 in English

According to Figure 3.4, there is no detection of an unusual (low) score of student H from Assessment 3 to Assessment 9. After Assessment 10, student H has suddenly departed from the normal stable scores and has only obtained 40. For student J, there is a detection of the unusual (low) score after the student completes Assessment 3 and it happens again at Assessment 10. The unusual (low) score of $Y_{J,1,3}$ is 50, whereas the unusual (low) score $Y_{J,1,10}$ is 40. Also, there is no detection of unusual (high) score for both students in mathematics from Assessment 3 to Assessment 10. Overall, the LCL and UCL for student H is predicted as higher than student J.

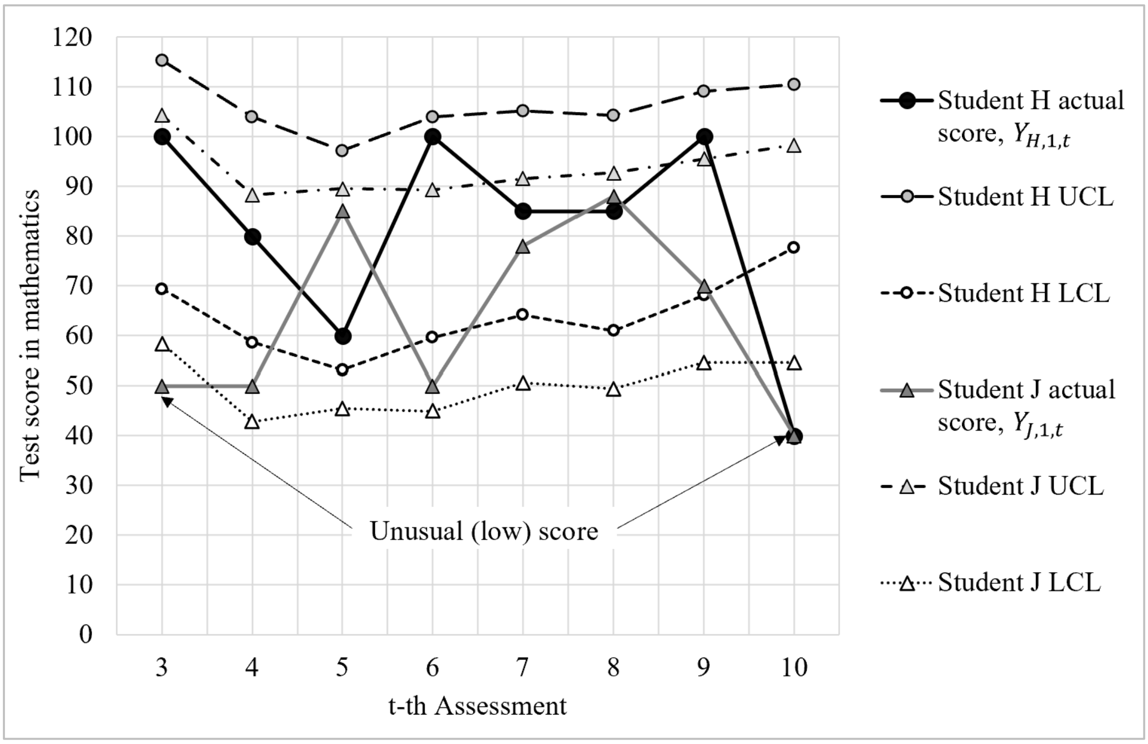


Figure 3. 4: Scores of student H and J from Assessment 3 to Assessment 10 for mathematics

3.4.3 Discussions

3.4.3.1 Interpretation of estimates of effects

Based on Figure 3.4 and from the analysis of predicted score $\hat{Y}_{H,j,10}$, among the 11 students, student H has average ability as indicated from Assessment 3 to Assessment 10 for all subjects in the classroom. The student's ability is represented by $\hat{\alpha}_H = 1.69$. The ability of student H is calculated based on equation (3.6), as an example of the computation for the estimate;

$$\hat{\alpha}_H = \left(\frac{Y_{H,1,1} + \dots + Y_{H,1,10} + Y_{H,2,1} + \dots + Y_{H,2,10} + Y_{H,3,1} + \dots + Y_{H,3,10}}{ST}\right) - \hat{\mu}$$
$$= \left(\frac{93 + \cdots + 40 + 85 + \cdots + 90 + 100 + \cdots + 80}{3 \times 10}\right) - 86.74 = 1.69.$$

For the difficulty of a subject $\hat{\beta}_1 = -1.16$, mathematics is estimated as more difficult than English $\hat{\beta}_2 = 0.24$ and science $\hat{\beta}_3 = 0.92$. Besides, Assessment 10 in mathematics $(\widehat{\beta\gamma})_{1,10} = 6.31$ is predicted as not a difficult assessment as compared to English, $(\widehat{\beta\gamma})_{2,10} = 2.18$ and science, $(\widehat{\beta\gamma})_{3,10} = -8.49$. It shows that because it is less difficult, students can obtain good scores in Assessment 10 of mathematics, compared to Assessment 10 for English and science. Considering specific ability of student H in mathematics $(\widehat{\alpha\beta})_{H,1} = -4.98$, this student has less ability in this subject compared to English $(\widehat{\alpha\beta})_{H,2} = 0.72$ and science $(\widehat{\alpha\beta})_{H,3} = 4.25$.

From the estimates, although student H's specific ability in mathematics is slightly less than in other subjects, it is almost equivalent to the average. Thus, it is informative for teacher to understand that the student has an interest in mathematics, which could be a motivation to produce a good score in Assessment 10. Also, since this assessment is not a difficult assessment compared to others, the teacher is able to recognize that it is not difficult to get high scores if student H seriously follows the evaluation criteria provided by the teacher. It is also a hint to the teacher that the content or questions can be understood by students in order to obtain their understanding of a particular topic.

3.4.3.2 Detection of unusual (low) scores

In Assessment 10, student H and student J have obtained unusual (low) scores $Y_{J,1,10}$ and $Y_{H,1,10}$ as 40, compared to other students who mostly scored perfect. This assessment is an individual activity IV and the students are individually assessed based on their specific abilities in the given task. The task needs to be submitted at the end of the classroom session. To ensure that student H and student J obtain high scores in the assessment, the teacher took the immediate remedial step of allowing the students to submit their tasks even after the deadline. The teacher verbally informed students about the submission several times. Unfortunately, the results were still the same.

As removing the unfavorable phenomena is important, inability to submit the task is recognized and it had disabled the teacher from providing scores for other criteria. As an additional remedy, the teacher may consider following up about the submission of the task by including parents in the reminding process. Another potential immediate remedy is to utilize other activities, equivalent to Assessment 10 (individual activity), to obtain the missing scores. The estimates of both students' abilities in mathematics and the difficulty level of Assessment 10 provide hints that the students can do the assessments with less guidance. Presumably, the students may not have submitted their assigned tasks due to private reasons. From a more detailed perspective of the estimates, teacher understood that student J has remarkably lower ability than student H. Thus, the approach to both students could be slightly different according to the ability of each student. For example, the reminder of task submission to both students could be the same, but teacher may assist student J more on how the task should be done to follow procedures and in terms of neatness in preparing task.

From a learning viewpoint, student H and student J are responsible to submit the tasks within the time frame, or at worst to still submit the task after the deadline. For the late submission, the students will not get the score for timely completion, but the teacher would still be able to provide scores for neatness and following procedures. By taking immediate remedy soon after Assessment 10, the unusual (low) scores for student H and student J in Assessment 10 can be stabilized for Assessment 11 onwards.

3.4.3.3 Information by least square estimates graph

This chapter provides an example of one of the least square estimates graphs obtained from the analysis. Figure 3.5 demonstrates the least square estimates graphs for the common causes of $\left(\widehat{\beta\gamma}\right)_{jt}$. Figure 3.5 shows the estimated difficulty level of T-th assessment for mathematics, English, and science, respectively. This information is beneficial for teachers to understand the difficulty level of each of the t-th assessment in a particular subject j and possibly help in improving the teaching quality. When teachers view this graph of estimates, it provides them indications to compare difficulty levels of each assessment along the semester both within a specific subject and between different subjects. Based on Figure 3.5, overall, all assessments are not remarkably far from the general mean. Thus, it could be interpreted that these assessments have almost similar difficulty levels.

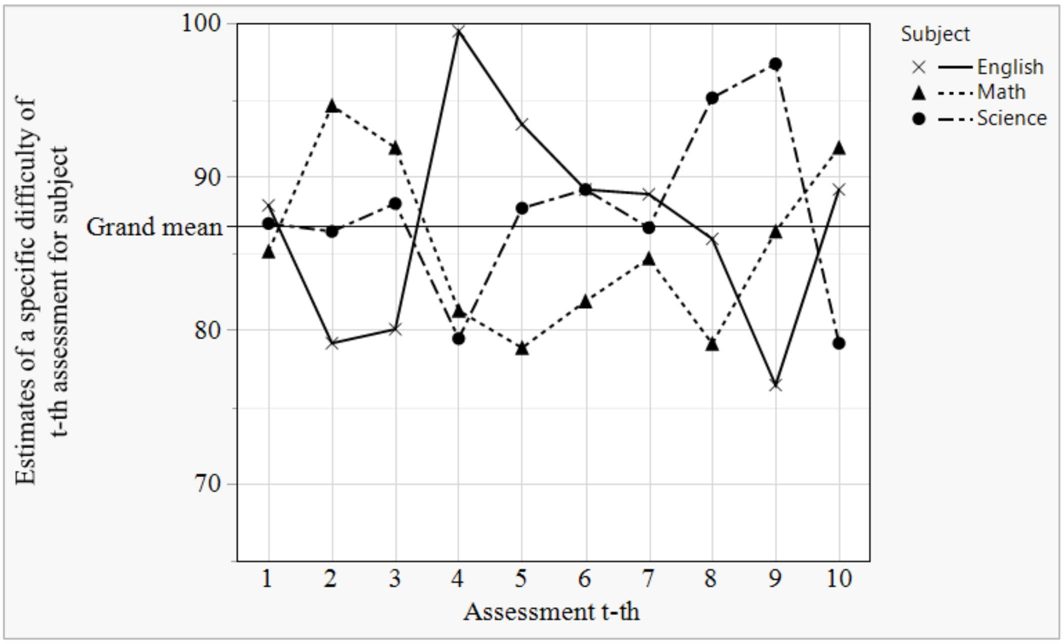


Figure 3. 5: Scores based on least-squared means for $\hat{Y}_{ijt}$ according to $\left(\widehat{\beta\gamma}\right)_{jt}$

3.4.3.4 At-risk scores and unusual (low) scores by 2×2 *matrix*

An at-risk score is a score which is likely to cause a student to fail in an assessment. For this case study, an at-risk score is represented by the threshold value, which is a cut-off score that the students should achieve in a subject. The threshold value decided by the school management for mathematics is 40. An at-risk score can also be represented by the fundamental ability of a student that is acquired while learning, listening, speaking, reading, writing and so forth. It helps to classify each student as a high performer or low performer.

For student H, this chapter confirmed that the unusual (low) score $Y_{H,1,10}$ is 40, which is equivalent to the threshold value of the at-risk score. The action of immediate remedies and root cause investigation can be interpreted based on the 2×2 matrix conditions of an at-risk score representing results, and an unusual (low) score representing the process of stability in education. This chapter maps the findings onto a 2×2 matrix; refer to Table 3.2. It is aligned to the four scenarios of occurrence of at-risk and abnormality (JSQC, 2014), each combination has specific actions that can be taken based on the existence of the at-risk score and unusual (low) score. The explanations of potential immediate remedies for case study (b) and (d) are described in 3.4.3.2 accordingly.

Table 3. 2: Appropriate actions based on findings in case study

	An unusual (low) score (abnormality) exists	An unusual (low) score (abnormality) does not exist
An at-risk score does not exist (it can be interpreted as a score that is above a threshold value or an ability for a student who is categorized as a high performer)	(b) Example: $Y_{J,1,3}$ is 50 Student J did not submit the task of making a video within the timeframe. As an immediate remedy, the teacher evaluated student J based on another equivalent activity in terms of the content to ensure that the student obtains a good score in Assessment 3. As an additional remedy, the teacher may further follow up on the Assessment 3 submission by the student to evaluate the performance. From a learning perspective, student J has to submit the tasks within the time frame. Even if the submission is overdue, it still has to be submitted for assessment. The teacher and student J should aim for (a) condition from the Assessment 4 onwards. In general, for the score in this quadrant, it could be predicted that the result in the next assessment would be poor if immediate remedies are not taken.	(a) Example: All scores other than $Y_{J,1,3}$, $Y_{J,1,10}$ and $Y_{H,1,10}$ Most of the students did well in Assessment 10 of mathematics. This can also portray that the students scored good consistently from Assessment 1 to 9. The school management should be glad that most student scores lie in this quadrant. Teachers should maintain the current way of teaching and assessing students' understanding for this cohort. Teachers should further monitor the results after conducting T-th assessment to ensure that this result is sustained. Students should also carry forward learning and reflecting the understanding into their assessment in the same manner for mathematics. The current efforts are sufficient to obtain good results; thus, they should keep up the work in each assessment. Schools should standardize the current teaching and learning method to maintain this result. This is the aim for process stabilization by the daily management concept in schools.
An at-risk score exists (it can be interpreted as a score that is below a threshold value or an ability for a student who is categorized as a low performer)	(d) Example: $Y_{J,1,10}$ and $Y_{H,1,10}$ are 40 Student H and J did not submit the tasks in the classroom as per the timeframe. As an immediate remedy, the teacher reminded both students to submit the tasks several times, but the students still failed to submit it. The scores obtained by students are based on their feedback of understanding in the classroom. As an additional remedy, the teacher can keep the parents of these students in loop of the reminding process to follow up on the task submission. Another potential remedy is to reassess students based on similar activities in the classroom at other times. As for the students, they should allocate time to complete the task appointed by their teacher, whether it is the same content or other content equivalent to Assessment 10.	(c) The example of case study is not available in this dataset Basically, as majority of the students in this cohort are average and high performers, the overall results have produced the general mean of more than 40 for Assessment 1 to 10. This is the reason for not having a case study for (c). Teachers and school management should be content for not having the exemplifying cases in this quadrant.

3.4.3.5 Comparison of findings with typical statistical methods

This chapter analyzes the data by using typical statistical quality control methods. Figure 3.6 illustrates the $\bar{X} - R$ chart where the subgroup consists of 11 students for mathematics from Assessment 1 to 10. Figure 3.7 utilizes the individual and moving range $(I - MR)$ chart to demonstrate the scores of student H and student J in mathematics from Assessment 1 to 10. These tools are presented in detail in standard textbooks like Montgomery (2008).

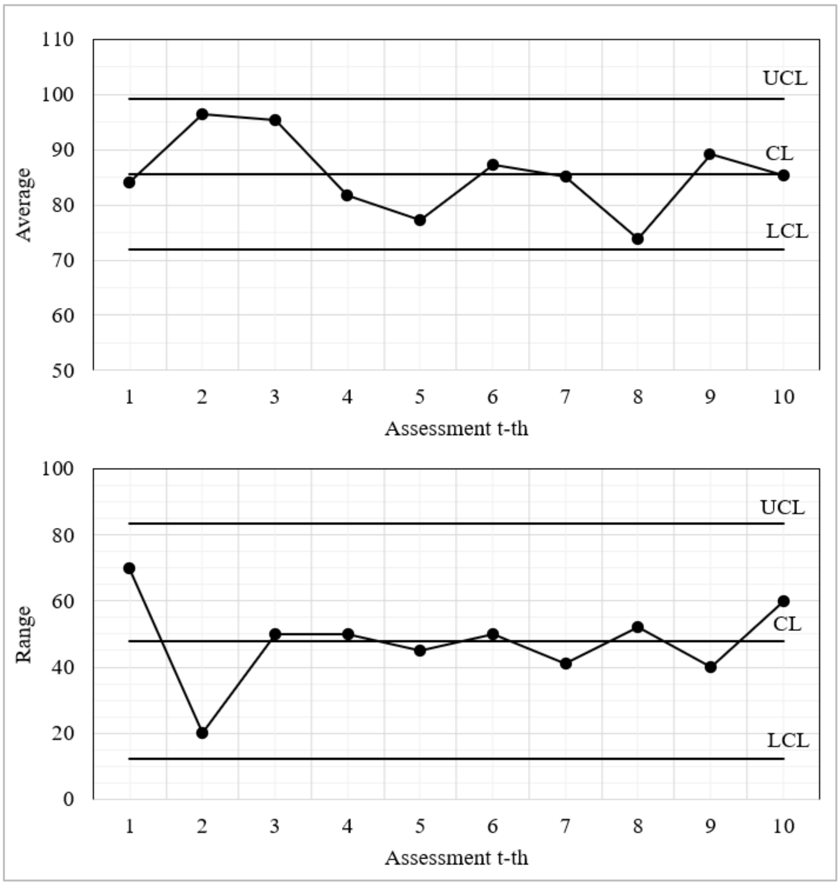
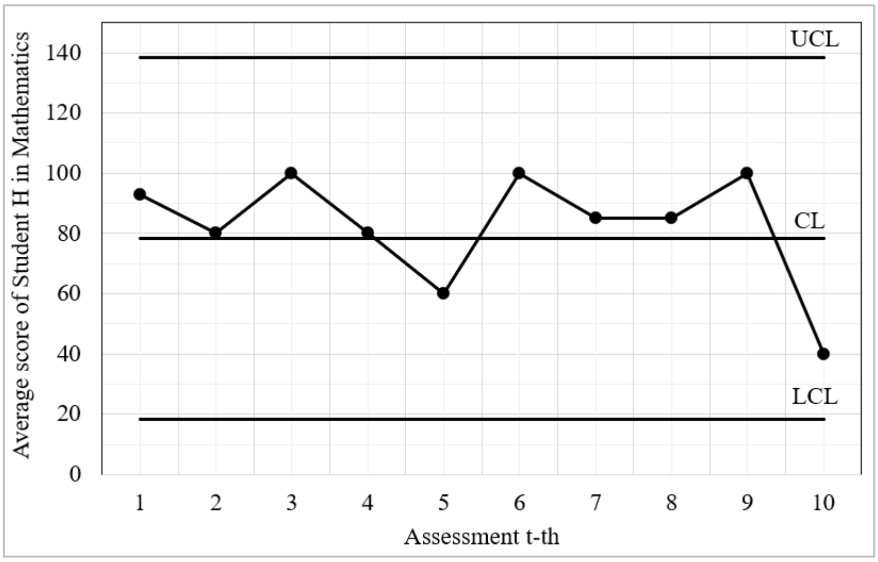


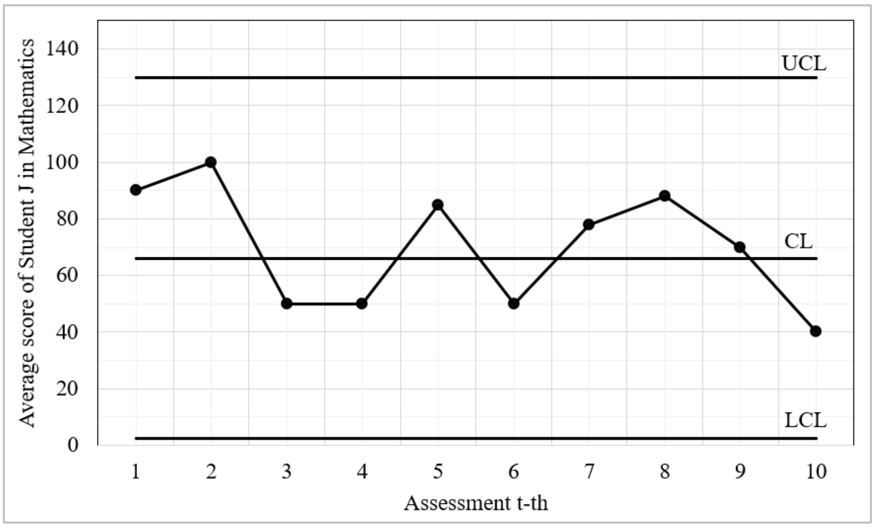
Figure 3. 6: $\bar{\boldsymbol{X}}$ and R chart of scores of 11 students in Assessment 1 to 10 for mathematics

In Figure 3.6, the LCL is consistent at 72.0 and 12.2 for $\bar{X}$ and R chart, respectively, because it is calculated based on the average value, although variation occurs in the subgroups data. There is no unusual (low) score detected for the dataset and all scores are within the controlled condition. Compared to the proposed model, almost similar information is demonstrated by $\left(\widehat{\beta\gamma}\right)_{1,t}$ in mathematics, refer to Figure 3.5. The information in Figure 3.6 is beneficial to analyze the existence of any unusual (low) score between the subgroups within 10 assessments. It is suitable to evaluate the variations of teaching quality in mathematics.

In Figure 3.7, the $I - MR$ chart of (a) and (b) establish the control limits for individual students. The LCL for student H and student J are consistent throughout all assessments in mathematics. The LCL of student H and student J are 18.3 and 2.26, respectively. There is no unusual score detected for the dataset and all scores are within the controlled condition. In Figure 3.7, the LCL is 3 times of the standard deviation. Unusual (low) scores are detected at $Y_{H,1,10}$ and $Y_{J,1,10}$ based on $2\hat{\sigma}$ of $LCL = \hat{Y}_{ijt} - k\hat{\sigma}$. This chapter also simulated the LCL for student H and student J based on $3\hat{\sigma}$ for a fair comparison with Figure 3.7. As a result, $Y_{H,1,10}$ and $Y_{J,1,10}$ are detected as an unusual (low) score.



(a) Student H



(b) Student J

Figure 3. 7: $\boldsymbol{I - MR}$ chart for student H (a) and student J (b) in mathematics for Assessment 1 until Assessment 10

Comparing Figure 3.7 with the proposed model, for example, the standard deviation of Figure 3.7(a) considers scores of student H in t-th assessments of mathematics only. Thus, it can be interpreted that the estimates of standard deviation include error variation and Assessment×Subject variation. From Figure 3.7(a), the estimate of standard deviation is 20.0, which includes the variation of error and Assessment×Subject. This estimate is large to detect an unusual score. Whereas for the proposed model, the $\hat{\sigma}$ only includes the error variation. The $\hat{\sigma}$ for the dataset in the case study is 10.94, which is close to the general mean. It illustrates the smaller range of controlled conditions or normal scores, therefore, more unusual scores can be detected. This chapter also analyzes the dataset based on time series models using an individual student score for one subject in several assessments. The detail analysis can be referred at Appendix 6. Based on selected time series models, the unusual (low) scores cannot be detected for student H at Assessment 10 in mathematics.

3.5 Conclusions

This chapter demonstrates a detection of unusual scores of individual students just after an assessment is carried out for immediate improvements. The availability and sharing of daily assessment activities data through the school monitoring system have made it possible to utilize the data for immediate remedy in the same academic year. A linear ANOVA model is developed based on a student's ability, a subject's difficulty level, a student's ability in a subject, and a difficulty level of an assessment in a subject as factor effects that affected student scores. Through least square estimates, a predicted score is generated. An unusual score is identified when the residual is below the pre-specified value as shown by LCL based on $k\hat{\sigma}$. The school considers $k = 2$ based on the balance of type I and type II errors, thus, $LCL = 2\hat{\sigma}$ is obtained separately for each student. A dataset of 11 students, 3 subjects, and 10 daily assessments in a semester is chosen for detailed analysis. The Assessment 10 of mathematics, English, and science are selected as examples because there are occurrences of unusual scores, compared to other t-th assessment. The unusual (low) scores of two students are detected in mathematics just after completing Assessment 10. There is no occurrence of an unusual (low) score for Assessment 10 in English and science, and these scores are categorized as normal. The detections of unusual (low) scores in Assessment 10 of mathematics are based on scores that suddenly deviate from other students in the same classroom. Besides, for these two students, unusual scores are further investigated for previous assessments in the same academic year. It is to see the occurrence of unusual (low) scores in t-th assessment before Assessment 10. On the other hands, there is no occurrence of unusual (high) score for all students in all subjects.

There are several practical implications of this chapter. Through the detection of unusual scores after each assessment is completed, particularly for unusual (low) score, immediate remedies can be taken to remove unfavorable phenomena to ensure student score is improved in the next assessment. This is according to the basic concept of daily management practice. The detection is visualized by using a control chart style to easily understand the problem. Also, by considering the factor effects of a student, subject, and assessment in a linear model, school management may get a quick indication of the individual student's abilities, difficulties of subjects, the ability of the student for a subject, and the difficulty level of assessments for a subject. Thereby, the detection can be linked to the problem, whether it is on the learning side or teaching side. It provides hints to teachers to take an immediate remedy for a specific student, regardless of their ability. The detection of unusual scores can be embedded in a school monitoring system for student progression in daily assessments. It encourages schools to further integrate all student scores in a centralized monitoring system for better decision making. The proposed model may also be utilized for academic subjects other than mathematics, English, and science. The condition of the data is by the assumption of it is normally distributed. Moreover, the model may use data based on similar fields, such as students' daily performances in non-academic courses, considering some adjustments. For example, those subjects that do not use a letter grading system shall adjust the scores in a numerical grading to fully utilize the model.

This chapter has some limitations. This chapter uses a dataset from only one school to validate the model. The proposed model is applicable only for scholastic subjects with a numerical grading system. Also, the mid-term and final term examinations are excluded from the dataset as these assessments have different distributions compared to daily assessments. Moreover, the t-th assessment recommended for analysis using this method is from Assessment 3 onwards to ensure more accuracy in estimates. When the number of assessments is becoming smaller, such as less than 6, the type II error will increase. It may not work well in terms of type II error compared to the case of 10 assessments. Furthermore, the number of students influences the factor effects of Student, Student×Subject, and error variance, which causes changes in DF. If there are more students in a cohort, the DF of error variance will be higher. The mean square error depends on the distribution of the data, which means that $\hat{\sigma}$

could be higher or lower than the value in the case study. Compared to the case study, if the $\hat{\sigma}$ is lower than 10.94, the LCL for individual students will be higher. In this chapter, an assumption is made that the dataset is normally distributed. The condition of distributions for future work data shall be closed to a normal distribution to ensure the type II error is not likely to occur. In future, it is recommended to carry out some simulation studies to examine the occurrence of type I and type II error to ensure the detection of unexpected score by the model is reflecting the truly critical or outstanding score obtained by students. Henceforward, the continuation of detection after assessments as part of students' performance monitoring system enhances the practice of daily management in teaching and learning and establishes a certain level of stability in school systems. It is expected that the model will help advance the field of knowledge in the specific area of process stabilization in schools and other educational institutions.

CHAPTER 4: Detecting unexpected scores of individual students in an examination based on past scores and current daily efforts

4.1 Introduction

Total quality management theories and concepts have been extensively applied in the field of education, including research on higher education (Sakthivel and Raju, 2006; Al-shafei *et al.*, 2015; Jasti *et al.*, 2022) but also in schools (Ghani and Pourrajab, 2014; Sfakianaki, 2019; Glaveli *et al*., 2022). The education sector is undergoing rapid changes to adapt to new challenges (Kwan, 1996), and the rising trends of globalization and internationalization have resulted in the diversification of society's educational needs. Under such circumstances, people demand higher quality and learning methods from educational institutions. In addition, educational institutions seek to improve the quality of education owing to moral, professional, competitive and accountability imperatives (Sallis, 2002).

The development of a monitoring system enables academics to observe the growth of individual students and monitor their learning progress longitudinally. These systems are designed to manage various parameters to improve the educational quality and encourage data-based decision-making by providing data for goal setting (Vlug, 1997; Schildkamp and Archer, 2017). Owing to the progress of information technology in schools, daily assessment data are available in centralized school monitoring systems. Alauddin and Yamada (2022) utilizes the students' daily assessment scores to predict the score of the latest daily assessment, which is an approach for daily management, especially the detection of an unusual state, to apply an immediate remedy and as a hint of good practices. An unusual score is a result that deviates from the normal score of an individual student. Particularly for unusual (low) score, the detection of an unusual (low) score alerts teachers and school management to apply immediate remedies to restore students to their normal academic level in the next daily assessment.

As the scores of previous examinations and daily assessments are both becoming available in centralized school monitoring systems, it is possible to utilize these data to analyze the detection of unexpected scores in an examination. In this chapter, it is aimed to predict the students' scores in a current examination by considering the past scores, current efforts in daily activities and trend in the current score. In this chapter, the past score is referred to the previous scores of the same subject. For example, the scores earned in mathematics examinations in past semesters can be used to predict the mathematical score in the current semester. A linear model is developed based on a combination of an ANOVA, a principal component analysis and a multiple regression analysis to predict the current examination score. The model is used to detect unexpected scores after the semester examination. The goal of detecting unexpected scores which is below the lower limit is to ensure immediate remedies to help students with critical scores. On the other hands, the detection beyond the upper limit is to indicate a good practice in learning and teaching based on a truly outstanding score. A case study is conducted using a dataset from an international school in Japan to validate the unexpected score detection method.

4.2 Literature review

4.2.1 Use of background data to predict scores

The prediction of students' academic performance has been researched to provide suggestions for educators to take proactive measures for timely intervention. In this chapter, the term background data refers to students' scores in the past, such as related pre-requisite courses, historical records of other subjects and attributes. The background data are the inputs to predict student scores. O'Connell *et al.* (2018) explores ten years of historical data of over 20,000 students of an introductory college-level algebra course. It intends to understand the relationship between course success and student performance in previous courses, student demographic background and time spent on coursework by using multiple regression analysis and principal component analyses. The findings demonstrate that the indicators of students' past performance and experiences, including the grade point average and the number of accumulated credit hours, best predict student success in the current course. Tomasevic *et al.* (2020) analyzes data comprising past examination or assessment scores, student engagement in activities and student demographic data, such as gender, age and highest education, to predict the final examination score. The student engagement data are related to search activity, the number of visits to learning sites, duration of learning, participation in related discussions, deck comments and the number of attempts to clear the examination.

A complete dataset based on a specific course, which includes related pre-requisite courses, can be used to predict the students' scores. Huang and Fang (2013) predicts student scores for a highly enrolled course using several predictive models. Data are gathered from students pursuing various majors enrolled in the course. The findings demonstrate that the selection of a predictive model depends on whether the instructor aims to improve the average performance of the classroom or the performance of individual students.

The scores of students can be used to analyze the differences among faculty members. For example, Otavová and Sýkorová (2016) analyzes the differences among the test scores obtained by students of mathematics courses from five faculties. The five levels of the main factors related to the faculty, two levels related to the semester, and their interactions are analyzed using a linear model. The results demonstrate that the students taught by different faculties perform differently in the mathematics course, and these differences persist throughout the semester.

A series of background data, such as scores from high schools and standardized examinations, can be utilized to forecast students' academic performances in college. Abuqaaud and Bou Nassif (2021) proposes several regression models to evaluate the relative influence of high school grades and standard examinations in forecasting the students' academic performance in college. The factors for this analysis include gender, college selection, high school average, high school mathematics score, high school English score and high school Arabic score. The findings demonstrate that the best predictors of the cumulative grade point average can be identified using statistical analysis.

In addition to examination scores, evaluations of homework, assignments and exercises can be utilized to predict student scores. Yang *et al.* (2018) improves the prediction accuracy of a model for predicting at-risk students by combining multiple regression and principal component analyses to conduct necessary interventions. Data pertaining to video-viewing behavior, exercise grades, homework completion and quiz grades are obtained to predict scores for a blended calculus course. Six components are extracted through a principal component analysis, and the predictive performance of the regression is improved using these components. The model is associated with designed learning activities for the course; thus, it does not apply to other courses with different attributes and learning activities.

The influence of the learning activities on the final examination is also examined. Chang and Wimmers (2017) explores the individual and combined effects of weekly assessments and practice exams on medical students' final examination performance. The weekly assessments, which are completed at the students' discretion, contained the material covered throughout the week's session. Based on past years' weekly assessment scores, the study suggests that this assessment is a useful approach for evaluating students' growth, although it does not directly contribute to a final grade. Lu *et al*. (2018) predicts students' final academic performance using variables from a blended course consisting of data from well-designed online and traditional learning activities, such as student behaviors in video viewing, out-of-class practice, and quiz scores. The results demonstrate that a blended dataset combining online and traditional critical factors has the highest predictive performance compared to individual data sets.

Previous literature has utilized the background data of students instead of using past scores on the same subject for which the future score is to be predicted. The examination scores stored in past years are employed as the input factor to predict student scores because these data are commonly accessible in educational institutions' data management systems. In addition, instead of examination results, assessments for specially designed learning activities from previous years are also being studied to predict the final examination scores. This chapter uses the past scores for the same subject, and several subjects with the same data structure are considered. In addition to past scores, current daily efforts in the current semester are used to predict the current examination score. Generally, data from current daily assessments are often kept by individual teachers for self-monitoring and are not systematically shared with others. Due to advancements in information technology in schools, daily assessment data for the current semester are stored and shared in the centralized monitoring system at certain schools. This type of assessment offers more recent data that reflect students' immediate understanding for higher accuracy in anticipating the current performance of an individual student. In addition, an analysis of the current score trend from the current examination is included to reflect individual factors, such as students and subjects, on the current examination score. This chapter hypothesizes that a combination of past scores, current efforts in daily activities, and trend in the current score could predict the current examination score with high accuracy. The prediction is performed for several subjects simultaneously, considering the subjects with the same data structure.

4.2.2 Management based on student scores

Formative assessments have been used in the learning process to determine a student's capability to master the subject and the resources required to succeed. Summative assessments are used to determine the extent to which learning has occurred. An advantage of these evaluations is that school administrators and teachers may examine the scores to determine whether students are on the path to attaining their learning objectives (Dixson and Worrell, 2016; Dolin *et al.*, 2018).

The teaching and learning processes are performed daily; they must be handled systematically through standardization, and a monitoring system capable of offering timely feedback to school management should be applied. A standard curriculum works as a set of guidelines for teachers to prescribe topics that need to be taught, as well as the contents, learning methods and materials. Following these requirements, teaching and learning processes are essential in daily management practice and total quality management of learning. Japanese Society of Quality Control (JSQC) describes daily management as activities required to effectively achieve the organizational objectives with regard to the job performed by each unit in the organization (JSQC, 2014). It is a system that documents the processes related to a particular job, as well as plans for managing and improving the operation (Cassidy, 1996), and daily management is practiced in the industrial (Ohno and Bodek, 1988; Kennedy, 2019) and service sectors (Biskupska and Chandima Ratnayake, 2019; Ontengco, 2019; Wang *et al.*, 2019; Tresh *et al.*, 2020).

Daily management involves various operations, including the management of abnormalities, to ensure process stabilization for achieving the target. Unusual process variations in educational institution courses are detected by analyzing students' scores. The findings are useful for evaluating course performance and variations between groups for feedback on the teaching approach. Hanna *et al.* (2012) investigates the specific causes of variations in the course passing rate to improve student performance in future semesters. A p-chart is used to calculate the average proportion of students who received a C or better grade in five courses over nine years. In the p-chart for each course, control limits ranging from σ to 3σ are drawn to illustrate the statistical variations that can be expected by the institution under normal circumstances. Both the positive and negative control limits are used to represent the course circumstances. Similarly, Cervetti *et al.* (2012) establishes a performance control chart to track the class performance variance in the positive or negative direction. In the chart, the point near the upper control limit provides the basis for setting the expectation and can be a useful resource for other teachers. The point near the lower control limit indicates where modifications might be needed in the course.

Additionally, Alauddin and Yamada (2022) utilizes daily assessment scores based on the structure of three factors: student, subject and assessment period. Daily assessment data are analyzed using a three-factor ANOVA model, and the detection of unusual scores in the latest assessment is based on previous daily assessments in the same academic year.

4.3 A model for detection of unexpected scores

4.3.1 Description of data structure

The purpose of using the model is to detect unexpected scores of individual students after examinations are conducted in the current semester. Past examination scores and daily efforts in the assessments in the current semester are utilized for detection. The application of advanced information technology systems in schools enables the collection of various assessment data of students at an individual or group level, data on subjects, from the schools' centralized monitoring systems. The unexpected score can be interpreted based on two conditions: i) unexpected (low) score to represent a truly critical score, and ii) unexpected (high) score to describe a truly outstanding score obtained by students.

Figure 4.1 illustrates the data structure for past examinations, current daily assessments and the current semester for utilizing the model developed in this chapter. Two factors are considered: student i's score in subject j in past examinations and daily assessments over different semesters. The examinations and assessments are conducted on three academic subjects – English, mathematics and science – and this chapter uses past examination scores, such as semester 1 and semester 2 in Figure 4.1. Furthermore, the current daily assessment should be regarded as the current effort for daily activities and attendance. The daily assessment is performed through numerical grading, ranging from 0 to 100. This includes monthly assessments of individual activities, group activities, notebook maintenance, mini tests and quizzes. In addition, this chapter uses the score obtained in the semester examination conducted following a series of daily assessments in the current semester. The data distribution of the examinations may differ from daily assessments owing to their purpose and implementation.

Subject: Science ($j = 3$)									
Student (i)	Past semesters		Current semester						
	Semester 1 examination	Semester 2 examination	Daily assessment						Semester examination
			1	2	3	4	5	6	
A									
B									

Subject: Mathematics ($j = 2$)									
Student (i)	Past semesters		Current semester						
	Semester 1 examination	Semester 2 examination	Daily assessment						Semester examination
			1	2	3	4	5	6	
A									
B									

Subject: English ($j = 1$)									
Student (i)	Past semesters		Current semester						
	Semester 1 examination	Semester 2 examination	Daily assessment						Semester examination
			1	2	3	4	5	6	
A									
B									
·	·	·	·	·	·	·	·	·	·
·	·	·	·	·	·	·	·	·	·
·	·	·	·	·	·	·	·	·	·
N									

Figure 4. 1: Outline of the dataset for past examinations and current semester for application to the proposed model

4.3.2 A method for detection of unexpected scores

1) Overall model

To take immediate action for a student whose score is unexpected in the current examination, this chapter assumes a model based on the current examination score. The basis of the model for the detection of an unexpected score is that the current examination score $Y_{(c)}$ can be described as the sum of the function of the past semester's score; $f_{(p)}$, function of the daily efforts in the current semester; $g_{(c)}$, and trend in the current examination score; $h_{(c)}$ such that:

$$Y_{(c)} = f_{(p)} + g_{(c)} + h_{(c)} + \varepsilon. \qquad (4.1)$$

In equation (4.1), the term ε denotes an error with a variance of σ^2 and p and c denote the past and current, respectively.

Considering the data in Figure 4.1, let $Y_{(c)ij}$ be the examination score in the current semester of student i for subject j, where $i = 1, \dots, n$ and $j = 1, \dots, S$. In the example illustrated in Figure 4.1, $i = 1, \dots, 14$ denote students A, B, …, N. In addition, $j = 1, 2, 3$ denote English, mathematics and science, respectively. Functions $f_{(p)}, g_{(c)}$, and $h_{(c)}$ are determined by applying the data structure illustrated in Figure 4.1. Specific functional forms of $f_{(p)}, g_{(c)}$ and $h_{(c)}$ are presented below. The score $Y_{(c)ij}$ is predicted as $\hat{Y}_{(c)ij}$ by applying a multiple regression analysis:

$$\hat{Y}_{(c)ij} = \hat{f}_{(p)ij} + \hat{g}_{(c)ij} + \hat{h}_{(c)ij}. \qquad (4.2)$$

In addition, the estimate of the standard deviation of the error; $\hat{\sigma}$ is derived using the standard deviation of the residuals; $Y_{(c)ij} - \hat{Y}_{(c)ij}$. Finally, the score $Y_{(c)ij}$ is detected as unexpected (low) if it is below the Lower Control Limit (LCL) $\hat{Y}_{(c)ij} - q\hat{\sigma}$, that is

$$Y_{(c)ij} < \hat{Y}_{(c)ij} - q\hat{\sigma}.$$

For the unexpected (high) score, it is detected when the residual is beyond the Upper Control Limit (UCL) $\hat{Y}_{(c)ij} + q\hat{\sigma}$, that is

$$Y_{(c)ij} > \hat{Y}_{(c)ij} + q\hat{\sigma}.$$

The coefficient $q > 0$ is determined using type I and type II errors. A type I error implies the detection of a regular score as an unexpected score, whereas a type II error implies the non-detection of a truly unexpected score. In general, a large q value provides fewer detections, fewer type I errors and more type II errors. In contrast, a small q value provides more detections, more type I errors, and fewer type II errors. A suggestion of decision for coefficient k or q by school management can be referred at Appendix 7.

2) Summarization of past examination scores

This chapter summarizes the past examination scores of the i-th student, j-th subject and k-th past examination; $Y_{(p)ijk}$ for unexpected score detection by summarizing $\hat{f}_{(p)ij}$, where $i = 1, \dots, n; j = 1, \dots, S; k = 1, \dots, K$. In the example in Figure 4.1, two semester examinations are illustrated. Hence, $K = 2$. This chapter uses the term $\hat{f}_{(p)ij}$ to denote the past examination score in equation (4.2):

$$\hat{f}_{(p)ij} = u_1\hat{Y}_{(p)ij1} + \cdots + u_K\hat{Y}_{(p)ijK},$$

where weight $u_1, \ldots, u_K$ is determined through a multiple regression analysis based on equation (4.2). In addition, $\hat{Y}_{(p)ijk}$ $(k = 1, \ldots, K)$ indicates the score predicted by a two-factor ANOVA according to the following model:

$$Y_{(p)ijk} = \mu_{(p)} + \alpha_{(p)i} + \beta_{(p)j} + (\alpha\beta)_{(p)ij} + (\beta\gamma)_{(p)jk} + \varepsilon_{(p)ijk}.$$

In this model, K past examinations are considered replication. It represents a repetition of the number of times of past examinations to assess student understanding. The assumption is that the error variance of each assessment is independent of each other and not time-dependent.

In addition, $Y_{(p)ijk}$ is the score of student i in subject j on the k-th past semester examination. The terms $\mu_{(p)}, \alpha_{(p)i}, \beta_{(p)j}, (\alpha\beta)_{(p)ij}$ and $(\beta\gamma)_{(p)jk}$ are the general mean, ability of student i, difficulty of subject j, specific ability of student i in subject j, and specific difficulty of subject j in the k-th examination, respectively. The estimates of $\hat{\mu}_{(p)}, \hat{\alpha}_{(p)i}, \hat{\beta}_{(p)j}, (\widehat{\alpha\beta})_{(p)ij}$ and $(\widehat{\beta\gamma})_{(p)jk}$ are obtained using the least squares with the above ANOVA model. The prediction score is determined by:

$$\hat{Y}_{(p)ijk} = \hat{\mu}_{(p)} + \hat{\alpha}_{(p)i} + \hat{\beta}_{(p)j} + (\widehat{\alpha\beta})_{(p)ij} + (\widehat{\beta\gamma})_{(p)jk}.$$

The example in Figure 4.1 has two past examinations; therefore,

$$\hat{f}_{(p)ij} = u_1\hat{Y}_{(p)ij1} + u_2\hat{Y}_{(p)ij2}$$

is applied.

3) Summarization of daily efforts in current semester

As illustrated in Figure 4.1, the daily efforts for subject j is measured using m variables. To summarize m variables, a principal component analysis, which is widely used to summarize multiple variables, is applied for each subject. Let $Z_{(c)ij1}, \ldots, Z_{(c)ijM}$ be the component scores of student i in subject j for $1, \ldots, M$ principal components, respectively. The number of utilized principal components M is determined based on the eigenvalues. This chapter uses the term

$$\hat{g}_{(c)ij} = v_1Z_{(c)ij1} + \cdots + v_MZ_{(c)ijM},$$

where the weight $v_1, \ldots, v_M$ are determined through a multiple regression analysis based on equation (4.2).

4) Trend of the current score and detection of unexpected scores

The trend in the current score $\hat{h}_{(c)ij}$ is described through the model;

$$\hat{h}_{(c)ij} = \hat{\mu}_{(c)} + \hat{\alpha}_{(c)i} + \hat{\beta}_{(c)j},$$

where $\hat{\mu}_{(c)}, \hat{\alpha}_{(c)i}$ and $\hat{\beta}_{(c)j}$ denote the estimates of the general mean; $\mu_{(c)}$, ability of student i; $\alpha_{(c)i}$ and difficulty of subject j; $\beta_{(c)j}$. It indicates a simple structure in which the current trend is based on the sum of student ability and subject difficulty.

Based on the above terms of the past examination score; $\hat{f}_{(p)ij}$, current daily efforts; $\hat{g}_{(c)ij}$ and trend in the current score; $\hat{h}_{(c)ij}$, equation (4.2) can be described as follows:

$$\hat{Y}_{(c)ij} = u_1\hat{Y}_{(p)ij1} + \cdots + u_K\hat{Y}_{(p)ijK} + v_1Z_{(c)ij1} + \cdots + v_mZ_{(c)ijM} + \hat{\mu}_{(c)} + \hat{\alpha}_{(c)i} + \hat{\beta}_{(c)j}.$$

For the example in Figure 4.1, the model can be described as

$$\hat{Y}_{(c)ij} = u_1\hat{Y}_{(p)ij1} + u_2\hat{Y}_{(p)ij2} + v_1Z_{(c)ij1} + v_2Z_{(c)ij2} + \hat{\mu}_{(c)} + \hat{\alpha}_{(c)i} + \hat{\beta}_{(c)j}, \quad (4.3)$$

because it has two past semesters and two principal components.

The model has several assumptions to ensure that it is practical for analysis through student scores. It is assumed that the error variance is constant regardless of the scores in the assessments or examinations for each student and subject. In addition, it is assumed that the errors are normally distributed.

4.4 Analysis of data

4.4.1 Case study data and outline of analysis

Data are collected from the Global Indian International School (GIIS), Tokyo Campus. Secondary grade students containing 21 students are chosen for consideration. Based on the structure in Figure 4.1, each student's past scores, current daily assessments, and current examination scores are examined. Seven students had incomplete datasets for two years in a row because they had moved or switched to a different curriculum when moving to a higher grade. As such, 14 students ($n = 14$) are selected for detailed analysis; the students are denoted as A, B, …, N. In this case study, p and c denote the years 2020 and 2021, respectively. The past semester’s scores are those of semester 1 and semester 2 examinations in 2020. The current daily efforts are obtained from six daily assessments scores in the current semester of 2021. This chapter aims to detect unexpected scores during the examination in

2021 for immediate remediation.

To summarize the past semesters score, this chapter uses semester 1 and semester 2 examination scores in 2020. Three subjects ($S = 3$) are considered: English ($j = 1$), mathematics ($j = 2$), and science ($j = 3$).

To summarize daily efforts in the current semester, this chapter utilizes the daily assessment data of the semester in 2021 as the daily efforts in the current semester. Six daily assessments are conducted in this semester, denoted by $a_1, a_2, \dots, a_6$. These assessments can be described as follows: a_1 (notebook maintenance 1), a_2 (subject enrichment activity 1), a_3 (periodic test 1), a_4 (notebook maintenance 2), a_5 (subject enrichment activity 2), and a_6 (periodic test 2); thus, $m = 6$. Notebook maintenance is evaluated by the teacher when students submitted their notebooks. The notebooks must be submitted several times during the semester, and the score is assigned based on content, timely submission and neatness. The subject enrichment activity aims to strengthen the students' understanding of the subject through individual activities, group-based activities and quizzes. A periodic test is an evaluation using a mini-test. As these three types of daily assessments are evaluated twice in each semester, they are indicated as notebook maintenance 1, notebook maintenance 2 and so on.

Table 4. 1: Summary of analysis of variables

Student i	Subject j	$\hat{Y}_{(2020)ij1}$	$\hat{Y}_{(2020)ij2}$	$Z_{(2021)ij1}$	$Z_{(2021)ij2}$	$Y_{(2021)ij}$
A	English	93	82	0.76	−0.68	88
B	English	96	85	0.89	−0.32	85
•	•	•	•	•	•	•
•	•	•	•	•	•	•
•	•	•	•	•	•	•
N	English	74	62	−1.52	−0.88	66
A	Mathematics	83	88	1.96	0.68	94
B	Mathematics	91	97	1.86	0.69	99
•	•	•	•	•	•	•
•	•	•	•	•	•	•
•	•	•	•	•	•	•
N	Mathematics	34	39	−3.67	−0.37	66
A	Science	84	89	1.34	−0.11	93
B	Science	88	94	1.41	0.19	91
•	•	•	•	•	•	•
•	•	•	•	•	•	•
•	•	•	•	•	•	•
N	Science	49	54	-2.78	−1.27	56

Table 4.1 presents an outline of the data analysis process. Specifically, the data under $\hat{Y}_{(2020)ij1}$ and $\hat{Y}_{(2020)ij2}$ are obtained by summarizing past examination scores. In addition, $Z_{(2021)ij1}$ and $\mathrm{Z}_{(2021)\mathrm{ij}2}$ are obtained by summarizing daily efforts in the current semester. Finally, $Y_{(2021)ij}$ is predicted based on $\hat{Y}_{(2020)ij1}$, $\hat{Y}_{(2020)ij2}$, $Z_{(2021)ij1}$, $Z_{(2021)ij2}$, student and subject through the application of a multiple regression analysis. The LCL and UCL are calculated based on the predicted score and standard deviation.

4.4.2 Summarization of past examination scores

The ANOVA results based on the above two-factor model are presented in Table 4.2. Overall, Student, Subject, and Subject×Examination have statistically significant effects on student scores. $R^2 = 0.89$ suggests that 89% of the variations in students' scores can be explained by their general ability, the difficulty level of the subject, and the difficulty level of the examination.

Table 4. 2: Analysis of variances based on semester 1 and semester 2 examinations in 2020

Source	DF	Sum of Squares	Mean Square	F-ratio	Prob > F
Model	43	21274	505.21	7.55	<.0001
Error	40	2677	66.92		
Corrected Total	83	24401			

Notes: $R^2 = 0.89$, Adjusted $R^2 = 0.77$

Focusing on mathematics and student M's score in semester 1, $\hat{Y}_{(2020)M,2,1}$ is calculated. The parameter estimates for the main effects of student M's ability, mathematics as the subject, special ability of student M in mathematics, and difficulty level of the semester examination are obtained using the least squares method. The estimate of the general mean is 77.54. $\hat{Y}_{(2020)M,2,1}$ is calculated as:

$$\hat{Y}_{(2020)M,2,1} = \hat{\mu}_{(2020)} + \hat{\alpha}_{(2020)M} + \hat{\beta}_{(2020)2} + \widehat{(\alpha\beta)}_{(2020)M,2} + \widehat{(\beta\gamma)}_{(2020)2,1}$$
$$= 77.54 - 22.05 - 5.23 - 11.27 - 2.92 = 36.07.$$

The same calculation is performed for all students for all subjects, denoted as $\hat{Y}_{(2020)ij1}$. The predicted scores for all subjects in the semester 2 examination in 2020 is denoted as $\hat{Y}_{(2020)ij2}$. The results of the prediction scores are illustrated in the columns $\hat{Y}_{(2020)ij1}$, $\hat{Y}_{(2020)ij2}$ in Table 4.1.

4.4.3 Summarization of daily efforts in current semester

The principal component analysis is performed separately for each subject. This chapter considers an analysis of mathematics as an example. Based on the principal component analysis, the eigenvalue of the first principal component (Component 1) is 3.92, with a contribution rate of 65.29%. Furthermore, the eigenvalue of the second principal component (Component 2) is 1.06, and the contribution rate is 17.73%. As these two components explain 83.02% of the variation in the data, it is judged that the results can be sufficiently explained up to Component 2. Figure 4.2 illustrates the factor loading for Component 1 and 2.

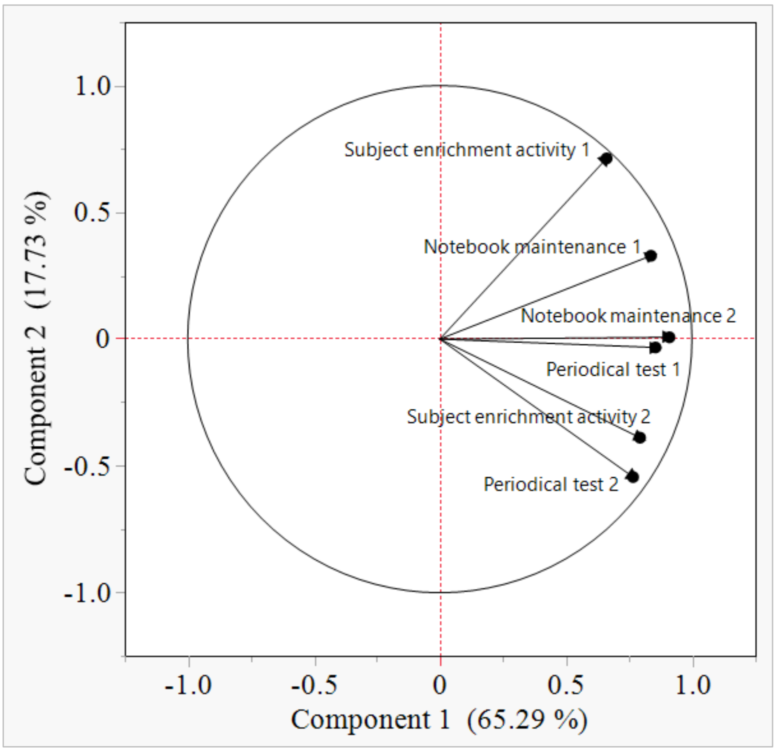


Figure 4. 2: Factor loading for Component 1 and 2

As illustrated in Figure 4.2, the factor loadings of Component 1 are positive. Based on the factor loadings, it can be said that Component 1 encompasses the overall efforts for the daily assessment, where positive indicates good overall effort and vice versa. This chapter interprets Component 1 as the students' overall effort in daily activities during the semester in 2021, denoted as $Z_{(2021)i,2,1}$. Regarding Component 2, the positive factor loadings are subject enrichment activity 1 and notebook maintenance 1, which are 0.71 and 0.33, respectively, whereas the negative factor loadings are periodic test 2 and subject enrichment activity 2, which are −0.54 and −0.39, respectively. This demonstrates that the first set of daily assessments in the current semester of 2021 has a positive factor loading, whereas the second set of daily assessments has a negative factor loading. Component 2 can be interpreted as the difference in effort between the first and second halves of the current semester and is denoted as $Z_{(2021)i,2,2}$. The results of the principal component score are illustrated in the columns of $Z_{(2021)i,2,1}, Z_{(2021)i,2,2}$ in Table 4.1.

4.4.4 Trend in the current score and detecting unexpected scores

A multiple regression analysis is applied to predict the current score based on the data in Table 4.1. The response variable is $Y_{(2021)ij}$. The input variables are student, subject, $\hat{Y}_{(2020)ij1}, \hat{Y}_{(2020)ij2}, Z_{(2021)i,2,1}$ and $Z_{(2021)i,2,2}$. The results are presented in Table 4.3. Based on the analysis, there is no evidence showing that the error variance is not constant. Detail of analysis can be referred at Appendix 5.

Table 4. 3: Analysis of variances based on multiple regression

Source	DF	Sum of Squares	Mean Square	F-ratio	Prob > F
Model	19	4724.69	248.67	5.59	0.0001
Error	22	978.09	44.46		
Corrected Total	41	5702.79			

Notes: $R^2 = 0.83$, Adjusted $R^2 = 0.68$

Based on equation (4.3), taking the example of student M and mathematics as the subject, the regression equation is obtained by performing a multiple regression as follows:

$$\hat{Y}_{(2021)M,2} = u_1\hat{Y}_{(2020)M,2,1} + u_2\hat{Y}_{(2020)M,2,2} + v_1 Z_{(2021)M,2,1} + v_2 Z_{(2021)M,2,2} + \hat{\mu}_{(2021)} + \hat{\alpha}_{(2021)M} + \hat{\beta}_{(2021)2}$$

$$\hat{Y}_{(2021)M,2} = -1.68\hat{Y}_{(2020)M,2,1} + 2.02\hat{Y}_{(2020)M,2,2} + 1.78Z_{(2021)M,2,1} + 4.63Z_{(2021)M,2,2} + 56.32 - 3.46 - 9.11$$

$$= 43.75 - 1.68\hat{Y}_{(2020)M,2,1} + 1.78Z_{(2021)M,2,1} + 2.02\hat{Y}_{(2020)M,2,2} + 4.63Z_{(2021)M,2,2}.$$

From the multiple regression, $\hat{\sigma}$ is 6.67, which is the estimate of the error variance for this analysis. Student M's ability in 2021 is $\hat{\alpha}_{(2021)M} = -3.46$, demonstrating that across all subjects, the ability of student M is considered lower than that of the other students. Moreover, science with $\hat{\beta}_3 = -9.36$ is more difficult than mathematics and English with $\hat{\beta}_1 = 18.47$.

Focusing on mathematics, $\hat{Y}_{(2021)i,2}$ is calculated for each student. To calculate the LCL, the school management selected $q = 2$ by considering the balance between type I and type II errors to indicate the LCL and UCL. The actual score of the current examination, predicted score, LCL and UCL for each student are plotted in Figure 4.3.

Figure 4.3 illustrates that an unexpected (low) score is detected for student M soon after completing the examination in 2021. The student obtained $Y_{(2021)M,2} = 50$, whereas the predicted score is $\hat{Y}_{(2021)M,2} = 65.5$. The scores of the other students are considered normal, and teachers could continue to monitor their progress through daily assessments in the next semester. Also, there is no detection of unexpected (high) score for all students in mathematics at current examination of 2021. As such, the description in this chapter is focusing on the occurrence of unexpected (low) score, while further explanation of detection of unusual (high) score that is similar to unexpected (high) is described in Appendix 8.

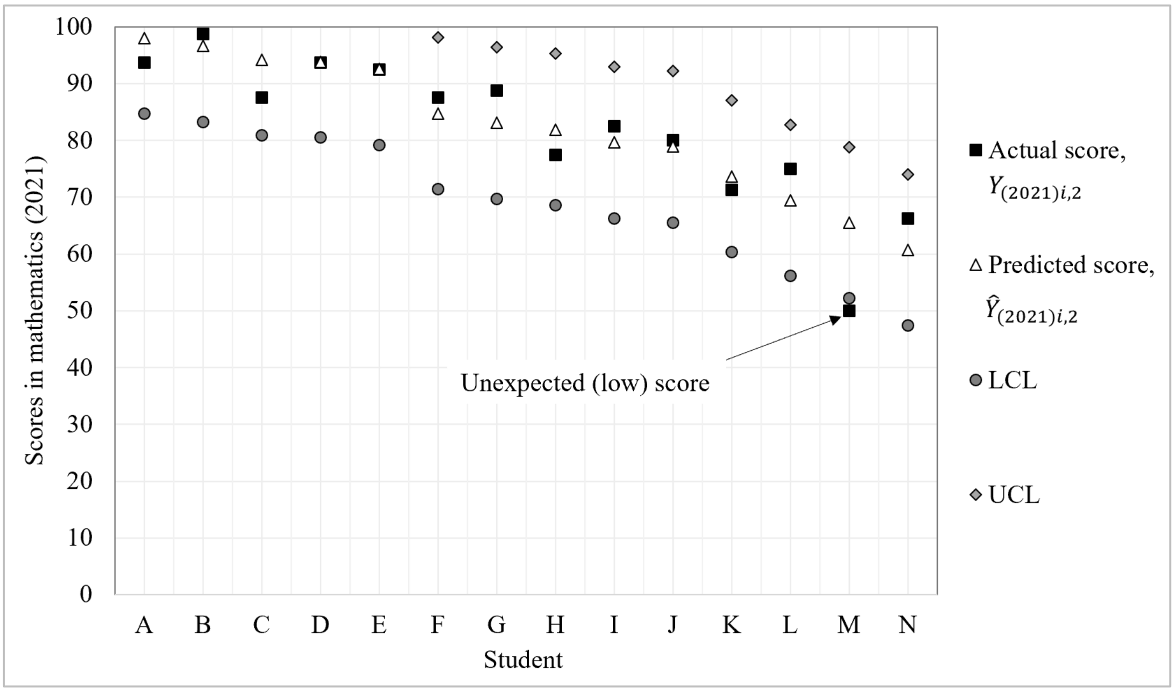


Figure 4. 3: Actual examination scores in 2021 against $\hat{\boldsymbol{Y}}_{(2021)i,2}$, with LCL and UCL of each student for mathematics

Student M has a lower ability in 2020, as indicated by $\hat{\alpha}_{(2020)M} = -22.05$, when compared with the other students. Moreover, this student exhibits lack of effort in the overall daily activities, for which $Z_{(2021)M,2,1} = -2.68$. Further, $Z_{(2021)M,2,2} = 0.56$ shows that student M performs better in the first half than in the second half of the current semester. Based on $LCL = 52.2$, the score of student M in mathematics in the examination in 2021 is detected as an unexpected (low) score. This detection can trigger teachers to take immediate remedial action for student M to ensure the score is restored to normal. In addition, student M must follow the instructions given by the teachers to improve the next assessment. As daily assessments are ongoing teaching and learning process, the immediate remedy taken after this examination will be reflected in the next set of daily assessments during the remaining semester period in 2021. This process is a continuous improvement cycle to ensure that student performance is normal.

In Figure 4.3, other than the unexpected (low) score, the predicted scores and the actual scores matched well as the high accuracy of prediction. Contrarily, the control limits based on $\pm 2\hat{\sigma}$ seems wide based on the difference between the two types of scores. There are three possibilities that might causing this phenomenon; existence of influential outlier, magnitude of error differed based on subjects, and distribution of residuals scores that by chance are near to control limit. Based on detailed analysis, the unexpected (low) score detected for student M is an outlier, but it is not quite likely influential to the model. For the magnitude of error differed based on subjects, based on unequal variance confirmation, there is insufficient evidence to claim that the variance between subjects is unequal. And this is supported by the assumption of error variance is constant for all subjects has been checked based on plot of residual against predicted score, stated in Figure XVI at Appendix 5. This phenomenon could be possible happen due to distribution of residuals scores for some students, such as in science, that are by chance near to UCL and LCL boundaries. The detailed analysis of three possibilities reasons is stated in Appendix 5.

Based on past scores and current daily efforts, student M is expected to obtain 65.5 in the current examination. As such, the unexpected score detected for student M is critical and requires an immediate remedy. The tendency to commit a type II error is higher, with fewer students in the case study. In this dataset, the usage of scores for 14 students in 3 subjects in two past semester scores, two extracted components from current daily efforts, and trend in the current examination through students and subjects is sufficient to predict the unexpected score of the current examination with an accuracy of 83%. The case study demonstrates a significant difference in student M's ability in the past and current semester, and it indicates to teachers why a student may perform differently over time. Analyses of current daily efforts and trend in current scores are important because it more accurately reflects the student's most recent performance. Particularly in daily assessments of mathematics, a student's efforts in the overall daily activities and a difference in terms of performance between the first and second halves of the current semester are factors that primarily influence the current examination scores. Thus, teachers should encourage students to actively participate in their current daily activities through their attendance, participation in individual or group activities, and task completion. Additionally, teachers can be alerted to the implementation of remedial actions for individual students, regardless of their ability.

For the model assumptions, this chapter plots the residuals and predicted scores to observe the pattern. The plots demonstrate no evidence that the error variance is not constant (see Figure XVI in Appendix 5). In addition, the Q-Q plot illustrated that the residual is approximately normally distributed (see Figure XVIII in Appendix 5). As these assumptions are fulfilled, the model is practically applied to the case study dataset to detect unexpected scores.

4.5 Comparison with other models based on previous literature

Using several models, this chapter analyzes the data discussed in the previous section. The comparison is made based on R^2, adjusted R^2, and $\hat{\sigma}$, as illustrated in Table 4.4. Model 1 is similar in approach to those used by Huang and Fang (2013) and Abuqaaud and Bou Nassif (2021), although these researchers use background data such as the scores in other subjects and in pre-requisite courses, whereas this chapter uses past examination scores. Model 2 is similar in approach to Otavová and Sýkorová (2016) and considers the students and some factors related to the courses attended. Models 3, 4 and 5 are the modified versions of the proposed model. This chapter simulates them based on a combination of two variables, namely, the semester examinations in 2020, daily efforts in 2021 and trend in the current score trend in 2021, to further observe the accuracy of prediction in comparison with that of the proposed model.

The results of Model 1 suggest that almost half of the variations in the student scores, which is 49%, can be explained by past scores. This chapter simulates Model 2 using the variables for current daily activities. The current score can be explained to the extent of 67%, based on student ability and subject difficulty. These values are lower than those of the proposed model, indicating that the combined use of past examinations, current effort and trend in the current score is effective in prediction compared with those of similar approaches proposed in the previous studies.

Table 4. 4: Comparison of several models in predicting student scores for current examination

	Semester examinations in 2020	Daily efforts in 2021	Student ability (α_i) and Subject difficulty (β_j)	R^2	Adjusted R^2	$\hat{\sigma}$
Proposed model	Included	Included	Included	0.83	0.68	6.67
Model 1	Included	Not included	Not included	0.49	0.47	8.56
Model 2	Not included	Not included	Included	0.67	0.49	8.44
Model 3	Included	Not included	Included	0.69	0.47	8.59
Model 4	Not included	Included	Included	0.81	0.67	6.70
Model 5	Included	Included	Not included	0.68	0.65	6.98

In addition, in Model 3, where this chapter adds the variables of past scores to Model 2, R^2 is further improved. In contrast, Model 5 utilizes only past scores and daily efforts in 2021 without considering the student's ability and difficulty level of the subject. Consequently, R^2 is slightly reduced compared to that of Model 3. It can be interpreted that although past scores and current daily efforts are available to predict student scores, the analysis still needs to consider the variations in the student ability and subjects to derive a more accurate model. When the analysis simulates a combination of the daily efforts in 2021, student ability, and subject difficulty, it suggests that 81% of the variations in the student score could be explained by these variables. This demonstrates that the daily efforts in 2021 is an important variable for improving the accuracy of predicting student score. Past scores are obtained over six months. Students may have changed their learning methods during the current semester. Therefore, the current daily efforts and trend in the current score most closely reflect student performance. By combining all three variables, a more accurate score prediction could be achieved for the current examination.

Figure 4.4 illustrates the detection of an unexpected (low) score by the proposed model and the models inspired by previous literature, namely Model 1 and Model 2. The score of $Y_{(2021)M,2} = 50$ for student M is identified as an unexpected (low) score by the proposed model because of $LCL = 52.2$. In contrast, this score is not identified as unexpected (low) by either Model 1 or Model 2. This difference is due to the prediction accuracy of the proposed model, as illustrated in Table 4.4.

Based on the comparison, it is clarified that the combinatorial usage of the current daily efforts, past scores and trend in the current score effectively detects unexpected scores. This combination predicts the current examination score with greater accuracy. This finding supports the hypothesis that the current examination score can be predicted with high accuracy using a combination of past scores, current efforts in daily activities, and trend in the current score. This demonstrates that instead of relying solely on past scores, school management may fully leverage the input from ongoing teaching and learning in the current semester, along with past results, for greater prediction accuracy. In addition, this result provides hints for school management to ensure the availability of both past scores and current daily assessments in a centralized monitoring system for further analysis to provide an early mitigation mechanism. This finding encourages teachers to ensure that students engage in current daily activities, as this significantly impacts how well they score on current examinations. Furthermore, students will be motivated to do better in their current daily activities and will be able to obtain help immediately if a problem arises. The detection of an unexpected (low) score alerts teachers that the unexpected (low) score is a truly critical score for the student and requires immediate remedial actions.

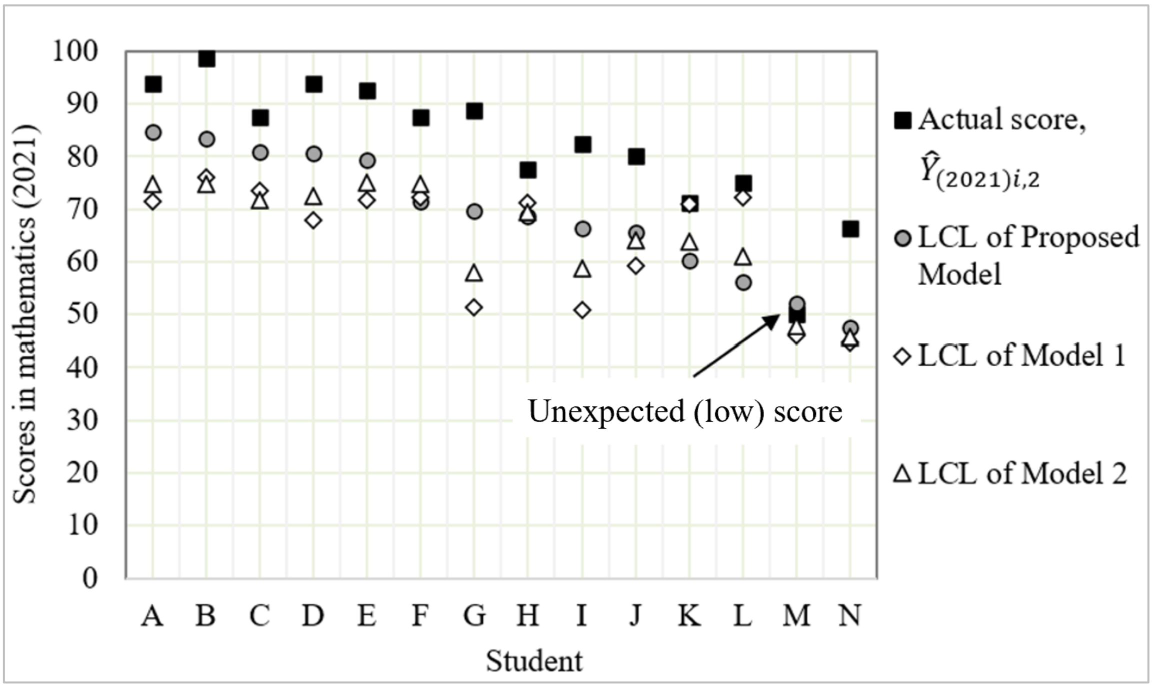


Figure 4. 4: Actual examination scores of each student in mathematics in 2021 along with LCL of the proposed model, model 1 and model 2

This chapter also analyzes the dataset based on time series models using past scores and current daily efforts. In this analysis, the current daily assessments are assumed and treated equally with past examination scores. All these assessments are used as past data. The detail analysis can be referred at Appendix 6. Based on selected time series models, the unexpected (low) scores cannot be detected for student M for current examination in mathematics.

4.6 Conclusion

4.6.1 Practical implications

This chapter proposes a model for detecting unexpected scores of individual students after an examination, with the objective of implementing immediate interventions. This chapter develops a detection method of unexpected score based on a combination of ANOVA, principal component analysis and multiple regression analysis. A case study is conducted on a dataset containing the data for 14 students, 3 subjects, past semester 1 and semester 2 examinations, 6 daily assessments in the current semester and the current semester examination. In the case study, two past scores are employed based on the main effects of student ability, subject difficulty, student's ability in a specific subject, and difficulty level of an examination using ANOVA. Schools may utilize more past examination scores and a larger amount of data, including data on students and subjects, to ensure that type II errors are unlikely to occur. The case study also demonstrates how the students' overall efforts in daily activities and the difference in efforts between the first and second halves of the current semester affected the current examination score in mathematics. Teachers may find it beneficial to perform a principal component analysis of the current daily assessments to identify the abilities that students can demonstrate in daily activities for each subject. For instance, students' ability to succeed in daily activities involving tests or their efforts to complete and submit notebooks may impact the results of the current examination, depending on the subjects. As such, teachers can promote greater learning through daily activities by emphasizing these qualities to stimulate students' interest in a particular subject. Additionally, to reflect the trend in current examination scores, additional factors such as student demographic data or the amount of time spent in a classroom may be considered in addition to students and subjects.

A combination of past scores, current efforts in daily activities, and trend in the current score can accurately predict the current examination score, which helps to detect students' unexpected scores accurately. Based on the case study, an unexpected (low) score is identified when the residual is below the LCL, based on $2\hat{\sigma}$. Taking mathematics as an example, students' scores are predicted immediately after the completion of the examination, and an unexpected (low) score is detected for a single student. The scores of the other students are categorized as normal. By detecting unexpected scores after the current examination, immediate remedies can be implemented to restore the student scores to normal in the daily assessments conducted in the next semester. This approach is in accordance with the fundamental practice of daily management as an important total quality management activity. The detection method for unexpected scores can be integrated with the school monitoring system to predict students' performance in the current examination by utilizing data from past examinations and daily assessments. Schools need an advanced and user-friendly information technology system to encourage teachers to store their current daily assessment data. The necessity to enter data on student scores for all subjects as soon as the assessment is completed and within the same timeframe may influence the work culture of teachers. This could be an initial step toward comprehensively integrating all student scores in formative and summative assessments in a centralized monitoring system for achieving educational excellence.

4.6.2 Limitations

This chapter uses a dataset collected from one school for model validation. In the future, the model should be implemented in multiple schools, with various types of assessments and quantities of data. The proposed model applies to subjects with a numerical grading system, typically for academic subjects such as English, mathematics, and science. In addition, the proposed model can be customized to meet the requirements of non-academic subjects such as sports and art activities by converting letter grading systems into numerical grading systems and standardizing the frequency of assessments. In addition, the selected subjects for the analysis shall have a similar frequency of assessments or examinations in the past and current years. The past scores, current daily assessments, and current examination scores shall be available for the same subject for each student. The proposed model is applied under the assumption of the availability of complete data, which requires the treatment of missing data through a suitable method, such as mean value imputation. Furthermore, the number of students influences the factor effects of student, student subject, and error variance, which can cause changes in the degree of freedom. The degree of freedom of error variance is higher for a greater number of students in the classroom. In future, it is recommended to carry out some simulation studies to examine the occurrence of type I and type II error to ensure the detection of unexpected score by the model is reflecting the truly critical or outstanding score obtained by students. Despite these limitations, it is hoped that the proposed model will contribute toward advancing current knowledge in stabilizing teaching and learning to achieve schools' targets.

CHAPTER 5: Concluding remarks

A TQM model based on Deming criteria is developed for the implementation of TQM in school education. In the B category, there are mainly eight criteria that can be as important TQM activities in school and become beneficial tools to overcome hurdles in educational practices to achieve school objectives. On top of that, the model substitutes the customer-oriented terminology with society-oriented school objectives and strategies to visualize a balanced stakeholder in education in the accomplishment of school objectives and linking them with results. By putting the proposed TQM model based on Deming criteria into practice, schools will be given the abilities and tools they need to overcome obstacles, improve student learning, and raise their level of excellence in the future. Through TQM implementation using the proposed model, there are numerous ways to improve students' learning and enhance the management of education. Aligned to the society-oriented model which locates student and teacher as a center of learning, improving the teaching and learning process is put a high weightage as the core process in education. Enhancing the grades of individual students through intervention is one of the ways to improve the quality of education.

Because of the advancement of IT in certain schools, on top of past examination scores, the student daily assessment data are also shared in the centralized monitoring system. It is made possible for data utilization considering students, subjects, assessments, and so forth for immediate remedy in the same academic year. For the detection of unusual or unexpected scores for individual students, in this research, two methods are being developed according to the type of data used for analysis and the purpose of detection. The approach is made for a daily basis assessment and covers until semester examination which is important for student grading. This research has revealed how to use data to achieve process stability in teaching and learning from the standpoint of daily management. It clarifies the expectations and decisions in processes under normal and unusual circumstances based on student scores. On top of semester examinations, the student's efforts in daily assessments are important for close monitoring and active participation to ensure the student results are in a normal range. In this research, the approaches can be applicable for the lower and upper side of the control limit. The lower control limit is an indicator to assist students whose score is truly critical and to ensure that no student fails or is in an at-risk condition. Furthermore, if the system triggers unusual (low) situations, teachers must intervene immediately within the same academic year to restore students' performance to normal. Whereas the upper control limit is a triggering point that a student has exceeded the expectation level, which may lead to a good practice for the individual student or in a classroom. The investigation of unusual (high) score occurrence may lead to findings of becoming a good learning and teaching way to achieve such outstanding result. When a student has exceeded the individual expected performance, it can boost up the spirit to keep up the current way and be motivated to perform better in next assessment. The concept of reducing standard deviation through several factor effects in daily assessment data may be possible to be applied not only in education field, but also in other sectors such as manufacturing and services.

The abundance of data accessible in school systems might be leveraged for greater data-driven decision-making by school management through scientific analysis. Schools must continue to improve their information technology systems to store and share more data within a school. Only when the information technology system advances, the advanced statistical study be possible. In education, there may be still a lack of understanding about how statistical control of processes can be utilized to establish an understanding of outcome ranges that can be expected under normal conditions and in building a baseline for continuous improvement. Statistical process control is ideal for capturing educational processes such as student retention, progression, and graduation. It provides a means to monitor the relationship between quality in education and academic progression to assure student learning through assessment processes. The correlation between daily assessments and semester examinations needs to be analyzed in detail to ensure that student participation in daily classes is positively and effectively affected by students' accumulated evaluation. The detection methods developed in this research focuses on individual students as the smallest entity in school hierarchy. The approach can be improvised and expanded to the higher level, such as teachers, schools, and systems in order to find benefits for each level in education.

Schools shall continuously use and analyze data from centralized monitoring systems to validate the courses that they offer and to assess the satisfaction of their services. Furthermore, educational management is important in the provision of educational services to society-oriented stakeholders. It must be ready to take remedial action if the demands of society-oriented stakeholders are not realized as expected. The TQM model based on Deming criteria helps in the evaluation of school ability acquisition in the management of educational quality to attain school excellence. This research focuses on practical implementation of managing daily activities in teaching and learning. For future studies, there are various areas can be selected from criteria in this TQM model, such as continuous improvements and new educational services in education process, which benefits schools. The statistical techniques and equivalent tools for analysis which is appropriate for schools to maintain and improve their processes and operations can be utilized. After all, the goal of TQM is not about conducting TQM activities, but to achieve school excellence and increasing educational quality. It is believed that the proposed model and the methods of detection of unusual or unexpected scores would be one of effective instruments to contribute to the advancement of knowledge in the important area of quality in schools and other educational institutions.

REFERENCES

Abuqaaud, K. A. and Bou Nassif, A. (2021) 'A statistical model for GPA estimation based on high school grades and standard exams', *2021 Innovation and New Trends in Engineering, Science and Technology Education Conference (IETSEC)*, IEEE, pp. 1–5. doi: 10.1109/IETSEC51476.2021.9440493.

Al-shafei, A. I., Bin Abdulrahman, K., Al-Qumaizi, K. I., and El-Mardi, A. S. (2015) 'Developing a generic model for total quality management in higher education in Saudi Arabia', *Medical Teacher*, 37(sup1), pp. S1–S4. doi: 10.3109/0142159X.2015.1006607.

Alagiah, R., Delaney, D. and McManus, L. (2001) 'Determining the relationship between attendance and performance in accounting education', *Asian Review of Accounting*, 9(1), pp. 128–138. doi: 10.1108/eb060739.

Alauddin, N. and Yamada, S. (2019) 'Overview of Deming criteria for total quality management conceptual framework design in education services', *Journal of Engineering and Science Research*, 3(5), pp. 12–20. doi: 10.26666/rmp.jesr.2019.5.3.

Alauddin, N. and Yamada, S. (2022) 'Detection of unusual scores of individual students for immediate remedy according to daily assessment data', *The TQM Journal*, pp. 1–19. doi: 10.1108/TQM-04-2022-0125.

Anastasiadou, S. D., Zirinoglou, P. A. and Florou, G. S. (2014) 'The European Foundation Quality Management evaluation of Greek primary and secondary education', *Procedia - Social and Behavioral Sciences*, 143, pp. 932–940. doi: 10.1016/j.sbspro.2014.07.530.

Arcaro, J. S. (1995) *Quality in education: an implementation handbook*, Delray Beach, Florida: St. Lucie Press.

Arribas Díaz, J. A. and Martínez-Mediano, C. (2018) 'The impact of ISO quality management systems on primary and secondary schools in Spain', *Quality Assurance in Education*, 26(1), pp. 2–24. doi: 10.1108/QAE-06-2016-0028.

Baeten, M., Struyven, K. and Dochy, F. (2013) 'Student-centred teaching methods: can they optimise students' approaches to learning in professional higher education?', *Studies in Educational Evaluation*, 39(1), pp. 14–22. doi: 10.1016/j.stueduc.2012.11.001.

Bayazit, O. (2003) 'Total quality management (TQM) practices in Turkish manufacturing organizations', *The TQM Magazine*, 15(5), pp. 345–350. doi: 10.1108/09544780310502435.

Beaver, W. (1994) 'Is TQM Appropriate for the Classroom?' *College Teaching*, *42*(3), pp. 111–114. https://doi.org/10.1080/87567555.1994.9926837.

Beshah, B. (2012) 'Students' performance evaluation using statistical quality control', *International Journal of Science and Advanced Technology*, 2, pp. 1–6.

Biskupska, D. and Chandima Ratnayake, R. M. (2019) 'On the need for effective Lean Daily Management in engineering design projects: development of a framework', *2019 International Conference on Industrial Engineering and Engineering Management (IEEM)*, pp. 789–794. doi: 10.1109/IEEM44572.2019.8978660.

Bouranta, N., Psomas, E. and Antony, J. (2021) 'Findings of quality management studies in primary and secondary education: a systematic literature review', *The TQM Journal*, 33(3), pp. 729-769. https://doi.org/10.1108/TQM-02-2020-0020.

Boyd, M. P. (2012) ‘Planning and realigning a lesson in response to student contributions’, *The Elementary School Journal*, 113(1), pp. 25–51. doi: 10.1086/665817.

Cassidy, M. P. (1996) ‘Streamlining TQM’, *The TQM Magazine*, 8(4), pp. 24–28. doi: 10.1108/09544789610125298.

Cervetti, M. J., Royne, M. B. and Shaffer, J. M. (2012) ‘The use of Performance Control Charts in business schools: a tool for assessing learning outcomes’, *Journal of Education for Business*, 87(4), pp. 247–252. doi: 10.1080/08832323.2011.592870.

Chang, E. K. and Wimmers, P. F. (2017) , ‘Effect of repeated/spaced formative assessments on medical school final exam performance’, *Health Professions Education*, 3(1), pp. 32–37. doi: 10.1016/j.hpe.2016.08.001.

Chapleo, C. and Simms, C. (2010) ‘Stakeholder analysis in higher education’, *Perspectives: Policy and Practice in Higher Education*, 14(1), pp. 12–20. doi: 10.1080/13603100903458034.

Chapman, R. and Al-Khawaldeh, K. (2002) ‘TQM and labour productivity in Jordanian industrial companies’, *The TQM Magazine*, 14(4), pp. 248–262. doi: 10.1108/09544780210429861.

Connors, C. B. (2021) ‘Summative and formative assessments: an educational polarity’, *Kappa Delta Pi Record*, 57(2), pp. 70–74. doi: 10.1080/00228958.2021.1890441.

Crawford, L. E. D. and Shutler, P. (1999) ‘Total quality management in education: problems and issues for the classroom teacher’, *The International Journal of Educational Management*, 13(2), pp. 67–72.

Creswell, J. W. (2014) *Research Design: Qualitative, Quantitative, and Mixed Methods Approach*, Thousand Oaks, California: SAGE Publications.

Dale, B. G., Bamford, J., Bamford, D., and van der Wiele, A. (2016) ‘Self-assessment, models and quality awards’, *Managing Quality: An Essential Guide and Resource Gateway*. Chichester, UK: Wiley, pp. 283–299. doi: 10.1002/9781119302735.ch12.

Dalin, P. (2005) *School Development: Theories and Strategies*, London: Continuum.

Dixson, D. D. and Worrell, F. C. (2016) ‘Formative and summative assessment in the classroom’, *Theory Into Practice*, 55(2), pp. 153–159. doi: 10.1080/00405841.2016.1148989.

Dolan, E. L., Elliott, S. L., Henderson, C., Curran-Everett, D., St. John, K., and Ortiz, P. A. (2018) ‘Evaluating Discipline-Based Education Research for promotion and tenure’, *Innovative Higher Education*, 43, pp. 31–39. doi: 10.1007/s10755-017-9406-y.

Dolin, J., Black, P., Harlen, W., and Tiberghien, A. (2018) ‘Exploring relationship between formative and summative assessment’, *Transforming Assessment*. Chichester, UK: Wiley, pp. 53–80. doi: 10.1002/9781119173489.ch29.

Donnelly, L. F. (2014) ‘Daily Management Systems in medicine’, *RadioGraphics*, 34(2), pp. 549–555. doi: 10.1148/rg.342130035.

EFQM (2021) *The EFQM-Hamdan Model*, *The European Foundation for Quality Management*. Available at: https://www.efqm.org/lens-series/education (Accessed: 17 November 2021).

Enquist, B., Johnson, M. and Rönnbäck, Å. (2015) 'The paradigm shift to Business Excellence 2.0', *International Journal of Quality and Service Sciences*, 7(2/3), pp. 321–333. doi: 10.1108/IJQSS-03-2015-0032.

Felder, R. M., and Brent, R. (2005) 'Understanding student differences', *Journal of Engineering Education*, *94*(1), pp. 57–72. https://doi.org/10.1002/j.2168-9830.2005.tb00829.x

Ghani, M. F. and Pourrajab, M. (2014) 'Sustainable education through implementation of Total Quality Management', *Global Business and Economics Research Journal*, 3(12), pp. 42–52.

Gillies, D. (2015) *Human Capital Theory in education*, *Encyclopedia of Educational Philosophy and Theory*, Springer Singapore. doi: 10.1007/978-981-287-532-7.

Glaveli, N., Vouzas, F. and Roumeliotou, M. (2022) 'The soft side of TQM and teachers job satisfaction: an empirical investigation in primary and secondary education', *The TQM Journal,* Vol. 34 No. 5, pp. 922-938. doi: 10.1108/TQM-11-2020-0269.

Goldberg, J. S. and Cole, B. R. (2002) 'Quality Management in education: building excellence and equity in student performance', *Quality Management Journal*, 9(4), pp. 8–22. doi: 10.1080/10686967.2002.11919033.

Hanna, M. D., Raichura, N. and Bernardes, E. (2012) 'Using statistical process control to enhance student progression', *Journal of Learning in Higher Education*, 8(1), pp. 71–82.

Hopkins, S., Walter, M. and Collins, S. (2017) 'Lean Daily Management: exploring the effectiveness in reducing producreturns and overnight shipment occurrences in a materials management department', *The Health Care Manager*, 36(3), pp. 267–272. doi: 10.1097/HCM.0000000000000173.

Huang, S. and Fang, N. (2013) 'Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models', *Computers* and *Education*, Elsevier Ltd, 61(1), pp. 133–145. doi: 10.1016/j.compedu.2012.08.015.

Ikemoto, G. S. and Marsh, J. A. (2007) 'Cutting through the "data-driven" mantra: different conceptions of data-driven decision making', *Yearbook of the National Society for the Study of Education*, 106(1), pp. 105–131. doi: 10.1111/j.1744-7984.2007.00099.x.

ISO (2015) *ISO 9001:2015 Quality Management Systems requirements*, *International Organization of Standardization*. Available at: https://www.iso.org/obp/ui/#iso:std:iso:9001:ed-5:v1:en (Accessed: 16 November 2021).

ISO (2019) *ISO 21001:2018 Educational organizations — Management systems for educational organizations — Requirements with guidance for use*, *International Organization of Standardization*. Available at: https://www.iso.org/obp/ui/#iso:std:iso:21001:ed-1:v1:en (Accessed: 17 November 2021).

Jasti, N. V. K., Venkateswaran, V., Kota, S., and Sangwan, K. S. (2022) 'A literature review on total quality management (models, frameworks, and tools and techniques) in higher education', *The TQM Journal,* Vol. 34 No. 5, pp. 1298-1319. doi: 10.1108/TQM-04-2021-0113.

Johnson, F. C. and Golomskiis, W. A. J. (1999) 'Quality concepts in education', *The TQM Magazine*, 11(6), pp. 467–473. doi: 10.1108/09544789910298852.

JSQC (2014) '*Guidelines for Daily Management (JSQC-Std 32-001(E): 2014)* ', The Japanese Society for Quality Control, Tokyo.

JSQC (2017) 'Guidelines for Policy Management (JSQC-Std 33-001 (E): 2017)', The Japanese Society for Quality Control, Tokyo.

JUSE (2021) *The application guide for the Deming Prize, the Deming Grand Prize*, Union of Japanese Scientists and Engineers, Tokyo. Available at: https://www.juse.or.jp/upload/files/DP_en_oubo2021v1.pdf (Accessed: 17 November 2021).

Kanji, G. K., Malek, A. and Tambi, B. A. (1999) 'Total quality management in UK higher education institutions', *Total Quality Management*, 10(1), pp. 129–153. doi: 10.1080/0954412998126.

Kaufman, P. and Bradbury, D. (1992) *Characteristics of at-risk students in NELS: 88*, Washington DC.

Kennedy, R. K. (2019) *Understanding, measuring, and improving Daily Management*, 1st edn., Boca Raton: Productivity Press. doi: 10.4324/9780429491672.

Khalaf, M. A. and Salem, T. S. M. (2018) 'The moderating effect of structural barriers on TQM-performance relationship in Egyptian service organizations', *International Journal of Quality and Service Sciences*, 10(4), pp. 349–365. doi: 10.1108/IJQSS-04-2017-0035.

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

Kumar, M. R. (2007) 'Comparison between DP and MBNQA: convergence and divergence over time', *The TQM Magazine*, 19(3), pp. 245–258. doi: 10.1108/09544780710745667.

Kumar, R., Garg, D. and Garg, T. K. (2011) 'TQM success factors in North Indian manufacturing and service industries', *The TQM Journal*, 23(1), pp. 36–46. doi: 10.1108/17542731111097470.

Kwan, P. Y. K. (1996) 'Application of total quality management in education: retrospect and prospect', *International Journal of Educational Management*, 10(5), pp. 25–35. doi: 10.1108/09513549610146114.

Lu, O., Huang, A., Huang, J., Lin, A., Ogata, H., and Yang, S. (2018), "Applying learning analytics for the early prediction of students' academic performance in blended learning", *Journal of Educational Technology & Society*, 2(2), pp. 220–232.

Mehra, S. and Rhee, M. (2004) 'Enhancing educational learning through some TQM principles', *International Journal of Quality* and *Reliability Management*, 21(8), pp. 801–816. doi: 10.1108/02656710410551719.

Mehta, A., and Degi, F. R. (2019) 'Total quality management implementation, and its barriers in education system', *International Scientific Journals*, *3*(1), pp. 37–43. https://doi.org/10.31219/osf.io/k26tq.

Militaru, M., Ungureanu, G. and (Crețu), A. Ș. C. (2013) 'The prospects of implementing the principles of total quality management (TQM) in education', *Procedia - Social and Behavioral Sciences*, 93, pp. 1138–1141. doi: 10.1016/j.sbspro.2013.10.003.

Miller, M. D., Linn, R. L. and Gronlund, N. E. (2009) *Measurement and Assessment in Teaching*, 10th edn., Pearson.

Montgomery, D. C. (2008) *Introduction to Statistical Quality Control*, 6th edn., Wiley.

Montgomery, D. C. (2012) *Design and analysis of experiments*, 8th edn., Wiley.

Montgomery, D. C., Jennings, C. L. and Kulahci, M. (2015) *Introduction to time series analysis and forecasting*, 2nd edn., New Jersey: Wiley. doi: 10.1201/9781315272443-7.

Montgomery, D. C., Peck, E. A. and Vining, G. G. (2012) *Introduction to Linear Regression Analysis*, 5th edn., New Jersey: Wiley.

Murgatroyd, S. (1992) 'A new frame for managing schools: total quality management (TQM)', *School Organisation*, 12(2), pp. 175–200. doi: 10.1080/0260136920120206.

Murgatroyd, S. J. and Morgan, C. (1993) *Total Quality Management and the School,* Buckingham: Open University Press.

Nakajo, T., Ando, Y., Munechika, M., Saito, T., Fukumaru, N., Yamada, S., and Mitsufuji, Y. (2019) *ISO 9004:2018 Explanation and application guidance*. 1st edn., Tokyo: Japanese Standard Association Solution Co.

Nakajo, T., Munechika, M. and Yamada, S. (2015) *ISO 9001:2015 Explanation of requirement*. 1st edn., Tokyo: Japanese Standard Association.

Nawelwa, J., Sichinsambwe, C. and Mwanza, B. G. (2015) 'An analysis of total quality management (TQM) practices in Zambian secondary schools', *The TQM Journal*, 27(6), pp. 716–731. doi: 10.1108/TQM-06-2015-0080.

NIST (2015) '2015-2016 Baldrige excellence framework (education) - A systems approach to improving your organization's performance', National Institute of Standards and Technology, Gaithersburg, MD.

O'Connell, K. A., Wostl, E., Crosslin, M., Berry, T. L., and Grover, J. P. (2018) 'Student ability best predicts final grade in a college Algebra course', *Journal of Learning Analytics*, 5(3), pp. 167–181. doi: 10.18608/jla.2018.53.11.

Ohno, T. (1995) *Workplace Management*. Cambridge, Massachusetts: Productivity Press.

Ohno, T. and Bodek, N. (1988) *Toyota Production System: beyond large-scale production*, New York: Productivity Press. doi: 10.4324/9780429273018.

Ontengco, J. B. (2019) 'Increasing referrals to a community paramedicine fall prevention program through implementation of a Daily Management System', *Journal of Trauma Nursing*, 26(1), pp. 50–58. doi: 10.1097/JTN.0000000000000415.

Otavová, M. and Sýkorová, I. (2016) 'Differences in results obtained by students of different faculties', *Journal on Efficiency and Responsibility in Education and Science*, 9(1), pp. 1–6. doi: 10.7160/eriesj.2016.090101.

Ozberk, O., Sharma, R.C., and Dagli, G. (2019) 'School teachers' and administrators' opinions about disability services, quality of schools, total quality management and quality tools', *International Journal of Disability, Development and Education*, 66(6), pp. 598–609. doi: https://doi.org/10.1080/1034912X.2019.164245.

Ruben, B. D., Russ, T., Smulowitz, S. M., and Connaughton, S. L. (2007) 'Evaluating the impact of organizational self-assessment in higher education: The Malcolm Baldrige/Excellence in Higher Education framework', *Leadership* and *Organization Development Journal*, 28(3), pp. 230–250. doi: 10.1108/01437730710739657.

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

Sakthivel, P. B. and Raju, R. (2006) ‘Conceptualizing total quality management in engineering education and developing a TQM educational excellence model’, *Total Quality Management* and *Business Excellence*, 17(7), pp. 913–934. doi: 10.1080/14783360600595476.

Sallis, E. (2002) *Total quality management in education*. 3rd edn., London: Kogan Page Ltd. doi: https://doi.org/10.4324/9780203417010.

Savic, M. (2006) ‘p-charts in the quality control of the grading process in the high education’, *Panoeconomicus*, 53(3), pp. 335–347. doi: 10.2298/PAN0603335S.

Scheerens, J. (2016) *Educational effectiveness and ineffectiveness*, Dordrecht: Springer. doi: 10.1007/978-94-017-7459-8.

Schildkamp, K. and Archer, E. (2017) ‘Feedback of monitoring data and its role in decision making at school and classroom level’, *Monitoring the quality of education in schools*, Rotterdam: Sense Publishers, pp. 11–24. doi: 10.1007/978-94-6300-453-4.

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

Shewhart, W. A. (1931) *Economic control of quality of manufactured product*, New York: D. Van Nostrand Company, Inc.

Sirvanci, M. B. (2004) ‘Critical issues for TQM implementation in higher education’ *The TQM Magazine*, *16*(6), pp. 382–386. https://doi.org/10.1108/09544780410563293.

Snyder, K. M. (2007) ‘The European Education Quality Benchmark System: helping teachers to work with information to sustain change’, *European Journal of Education*, 42(3), pp. 425–435. doi: https ://doi.org/10.1111/j.1465-3435.2007.00315.x.

Sohel-Uz-Zaman, A. S. M., and Anjalin, U. (2016) ‘Implementing total quality management in education: compatibility and challenges’, *Open Journal of Social Sciences*, *04*(11), pp. 207–217. https://doi.org/10.4236/jss.2016.411017.

Spanbauer, S. J. (1995) ‘Reactivating higher education with total quality management: using quality and productivity concepts, techniques and tools to improve higher education’, *Total Quality Management*, 6(5), pp. 519–538. doi: 10.1080/09544129550035189.

Syukur, M. (2021) ‘Roles of gender, study major, and origins in accounting learning: a case in Thailand’, *The International Journal of Management Education*, 19(3). doi: 10.1016/j.ijme.2021.100555.

Thandapani, D., Gopalakrishnan, K., Devadasan, S. R., Sreenivasa, C. G., and Murugesh, R. (2012) ‘Quality models in industrial and engineering educational scenarios: a view from literature’, *The TQM Journal*, 24(2), pp. 155–166. doi: 10.1108/17542731211215099.

Tomasevic, N., Gvozdenovic, N. and Vranes, S. (2020) ‘An overview and comparison of supervised data mining techniques for student exam performance prediction’, *Computers* and *Education*, 143. doi: 10.1016/j.compedu.2019.103676.

Töremen, F., Karakuş, M. and Yasan, T. (2009) ‘Total quality management practices in Turkish primary schools', *Quality Assurance in Education*, 17(1), pp. 30–44. doi: 10.1108/09684880910929917.

Tresh, A., Cohen, A. J., Mmonu, N. A., Berdy, S., Barnas, K., Krombach, J., and Breyer, B. N. (2020) 'Resident-driven holistic Lean Daily Management System to enhance care experience at a safety net hospital', *Urology*, 140, pp. 56–63. doi: 10.1016/j.urology.2020.01.038.

Tribus, M. (2005) 'TQM in education: the theory and how to put it to work', *Developing Quality Systems in Education*, 1st edn., Routledge, pp. 46–59.

UNESCO (2017) *Education transforms lives*, *United Nations Educational, Scientific and Cultural Organization*, Paris, France.

Venkatraman, S. (2007) 'A framework for implementing TQM in higher education programs', *Quality Assurance in Education*, 15(1), pp. 92–112. doi: 10.1108/09684880710723052.

Verma, P., Kumar, V., Mittal, A., Gupta, P., and Hsu, S. C. (2022) 'Addressing strategic human resource management practices for TQM: the case of an Indian tire manufacturing company', *The TQM Journal*, 34(1), pp. 29–69. doi: 10.1108/TQM-02-2021-0037.

Vlug, K. F. M. (1997) 'Because every pupil counts: the success of the pupil monitoring system in the Netherlands', *Education and Information Technologies*, 2(4), pp. 287–306. doi: https://doi.org/10.1023/A:1018629701040.

Vokurka, J. R., Standing, G. L. and Brazeal, J. (2000) 'A comparative analysis of national and regional quality awards', *Quality Progress*, pp. 41–49.

Wang, X., Fang, X., Jin, J., Bao, H., Wang, X., Gao, C., Hou, W. (2019) 'Real time monitoring of dams in dam daily management taking Xianlin Reservoir as an example', *IOP Conference Series: Earth and Environmental Science*, 304(2). doi: 10.1088/1755-1315/304/2/022042.

Whitley, S., Estes, D., Mckenzie, M., Salimi, K., Barada, B., and Merlino, J. (2015) 'Daily Management System to improve performance in a behavioral health primary care practice', *Journal of Family Medicine*, 2(6), pp. 1–8.

van der Wiele, T., Dale, B., Williams, R., Kolb, F., Moreno Luzon, D., Schmidt, A., and Wallace, M. (1995) 'State-of-the-art study on self-assessment', *The TQM Magazine*, 7(4), pp. 13–17. doi: 10.1108/09544789510092845.

Wirtz, A. L., Monsees, E. A., Gibbs, K. A., Myers, A. L., Burns, A. N., Lee, B. R., Goldman, J. L. (2021) 'Integration of a lean Daily Management System into an antimicrobial stewardship program', *Pediatric Quality & Safety*, 6(2), pp. 1–7. doi: 10.1097/pq9.0000000000000384.

Yang, S. J. H., Lu, O. H. T., Huang, A. Y. Q., Huang, J. C. H., Ogata, H., and Lin, A. J. Q. (2018) 'Predicting students' academic performance using multiple linear regression and principal component analysis', *Journal of Information Processing*, 26, pp. 170–176. doi: 10.2197/ipsjjip.26.170.

Yapa, S. (2012) 'Total quality management in Sri Lankan service organizations', *The TQM Journal*, 24(6), pp. 505–517. doi: 10.1108/17542731211270070.

Zarbo, R. J., Varney, R. C., Copeland, J. R., D'Angelo, R., and Sharma, G. (2015) 'Daily Management System of the Henry Ford production system', *American Journal of Clinical Pathology*, 144(1), pp. 122–136. doi: 10.1309/AJCPLQYMOFWU31CK.

ACKNOWLEDGEMENTS

First, thank you Allah, for His grace and mercy that finally I can complete my Ph.D. studies and the courage to make the hardest decision in my life, which is to choose to pursue my Ph.D. dream over my professional career.

This research has been conducted at Keio University, Graduate School of Science and Technology, Department of Open Systems Management. I would like to express my deep gratitude to my supervisor, Professor Shu Yamada for his patient guidance, useful critics, and precious sharing of his knowledge throughout the process. I am extremely lucky to have him as my supervisor and his existence is like a family member to me as I am alone in Japan, far from my family and relatives.

I would like to extend my special thanks to the Global Indian International School (GIIS), Tokyo Campus; the place where I collected my data and validated several new models for improvements of quality in education towards achieving school excellence. Not forgetting the important persons in my life; my lovely husband, Razak Bin Abu Bakar, and my dearest mother, Azizah Binti Mohamad, including all the family members, lab mates, and friends, without whom I was nothing. They not only assisted me morally but also extended their support emotionally when I was facing hard times. Although Covid-19 has slightly interrupted the research schedule, to some extent I managed to catch up and finally crossed the finishing line. Finally, to the rest who are directly or indirectly involved in the entire research activities, your support throughout this research is much appreciated.

Alauddin, Nursuhana binti
Yokohama, Japan
February 2023

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

Y_{ijt}	Score of student i for subject j on the t-th assessment
$\hat{Y}_{ijt}$	Predicted score of student i for subject j on the t-th assessment
$\mu, \hat{\mu}$	General mean and its estimate
$\alpha, \hat{\alpha}_i$	Ability of student i and its estimate
$\beta, \hat{\beta}_j$	Difficulty of subject j and its estimate
ε_{ijt}	Error term
$(\alpha\beta)_{ij}, \left(\widehat{\alpha\beta}\right)_{ij}$	An ability of student i in subject j and its estimate
$(\beta\gamma)_{jt}, \left(\widehat{\beta\gamma}\right)_{jt}$	A difficulty of t-th assessment for subject j and its estimate
$Y_{(c)ij}, \hat{Y}_{(c)ij}$	Student i obtains examination score for subject j in current examination and its estimates
$Y_{(p)ij1}, \hat{Y}_{(p)ij1}$	Student i obtains the past semester 1 examination score and its estimates
$Y_{(p)ij2}, \hat{Y}_{(p)ij2}$	Student i obtains the past semester 2 examination score and its estimates
$f_{(p)}$	Sum of the function of past semester score
$g_{(c)}$	Function of daily efforts in current semester
$h_{(c)}$	Trend of the current examination score
$\hat{f}_{(p)ij}$	Estimate of function of past score for student i in subject j
$\hat{g}_{(c)ij}$	Estimate of function of daily efforts in current semester for student i in subject j
$\hat{h}_{(c)ij}$	Estimate of trend of the current examination score student i in subject j
$Z_{(c)ij1}, Z_{(c)ij2}$	First two main components from principal component analysis
$\sigma, \hat{\sigma}$	Standard deviation of error and its estimate
$\hat{Y}_{ijt} - k\hat{\sigma}, \hat{Y}_{ijt} - q\hat{\sigma}$	Lower control limit (LCL)
$\hat{Y}_{ijt} + k\hat{\sigma}, \hat{Y}_{ijt} + q\hat{\sigma}$	Upper control limit (UCL)

APPENDICES

Appendix 1: A society-oriented model for school education: appropriate customers definition towards achieving school excellence

i) Philosophy of the model creation

In this research, there are two main factors should be considered comprehensively; flow of the value received from the school, and value provided to the school. The flows should be identified from each group to school, considering the potential parties or stakeholders with their activities. The development and learning processes to equip required knowledge and skills are fundamental elements for contemplation as well. The relationship between teacher and students in teaching and learning process is about these two parties in having frequent interaction and communication from initial stage, up to graduation day. Besides, process of communication between school and related parties is another criterion to be reflected and consideration of education timeline from student enrollment to school until they graduate is included in the model.

ii) Process to develop the model

Step 1: Create a list of potential interested society-oriented parties in categorization

Potential parties or stakeholders are listed up, considering all groups or people who has interest in school. Previous researchers have identified stakeholders, known as customers, in education institutions (Spanbauer, 1995; Kanji, Malek and Tambi, 1999; Militaru *et al*., 2013; Sallis, 2002; Chapleo and Simms, 2010), and it could be selected as input of potential groups or stakeholders. In determining what a stakeholder is, it is important to consider anyone who may fall into any of society-oriented categories having interest in school education. Then, these people or organizations are categorized under six categories and numbered based on its importance - 1) internal; 2) parties who are frequently involved with the school; 3) regulatory bodies; 4) surroundings; 5) immediate organizations after completion of education; and 6) other related parties. The parties are categorized according to its commonalities of interest to school, for example regulatory bodies consists of government, accreditation and regulatory bodies.

Step 2: Locate society-oriented categories according to its involvement at school

Locating the school internal as a starting point at left side of the model, locate all society-oriented categories according to educational timeline horizontally, from current to after graduation time. The categories are positioned according to vertical line showing the skills and knowledge utilization as individual basis on top and societal-basis at bottom side. As both skills and knowledge development are equally important, the arrow is pointing out in two-directional way. Moreover, the positioning of students and teacher is close to each other as they are the primary components in the society-oriented model for school education based on teaching and learning process. Both parties create a frequent communications and interactions with each other to ensure the learning process happens with appropriate assessment of understanding.

Step 3: Describe the provision of values from school to society and vice-versa in a table

The value from school to society and vice-versa with each category indicates the provision of values, whether tangible or intangible. The detailed description is shown in Table I. Tangible value can be classified as those values that can be touched, seen, and normally known as physical value, such as monetary and facilities. Intangible value is the opposite of tangible ones and it cannot be touched and seen, for example knowledge and supports. When clarifying the value, as priority, learning process is considered as a main idea to determine appropriate values to encourage this process to happen in school. As education is mainly about main communication between teacher and students, these are the main parties with more frequency interactions in provision of learning process. For students graduate from the school, some of them will proceed to higher educational institution and there are students who will immediately be hired by companies. Eventually, all of them will go back to community to contribute in improvement of society through skills and knowledge that they obtain from educational institutions.

iii) Model description

Figure 2.2 illustrates the society-oriented model for school education with its description in terms of provision of values between school and society in Table I. The main focus of society-oriented in the model is seen according to relationship between teacher and students. As education is primarily about teacher-student communication, especially within the daily teaching and learning process of transferring knowledge. As such, these are the parties with the most frequent interactions in the provision-of-learning process. Student is taught by teacher in multiple courses based on curriculum, syllabus, and subjects. The learning process starts from the time student enroll at school until completion of school. There are numerous parties involving in the process of learning indirectly and generally these people communicate more often with school staffs and top management to provide a comfortable, conducive and safety environment for learning, and at the same time secure the quality of education and operations.

Table I: Description of society-oriented model in terms of provision of values between school and society

Category	Society-oriented parties	Provision of value to school	Provision of value from school
1. Internal	Teacher	1. Frequent communication with students to provide learning 2. Loyalty and retention in school 3. Participation in achieving school mission	1. Hired as employees 2. Salaries (including bonus, welfare) 3. Conducive and safety facilities for teaching and working 4. Ability development for teaching
	Staff	1. Supporting learning of student 2. Loyalty and retention in school 3. Participation in achieving school mission	1. Hired as employees 2. Salaries (including bonus, welfare) 3. Conducive and safety facilities for teaching and working 4. Ability development for working
	Top management	1. Making decision for school mission, strategies 2. Ensure commitment of staffs to work 3. Overall performance monitoring	1. Hired as management level 2. Salaries (including bonus, welfare)
2. Parties who frequently involve with school	Student	1. Enrollment of self 2. Commitment, participation in learning	1. Provision of education services (academic, skills and so forth) 2. Facilities for learning 3. Conducive and safety learning place
	Parent	1. Enrollment of children 2. Tuition fee	1. Information of students' progress 2. Improve family quality of living, societal status through educated children

3. Regulatory bodies	Government	1. Ensure to regulate the learning opportunities 2. Government funds 3. Tax paid by tax payers 4. Hiring students as government employees	1. Fulfilment of laws and regulations 2. Required knowledge (skills, knowledge) for employment
	Accreditation bodies for education program	Secure the education quality	Assurance of education through accreditations
	General regulatory bodies	Assure the operation according to regulation standard to secure education quality	Assurance of quality in operation
4. Surrounding parties	Neighbor hood	1. Potential enrollment 2. Good ambient creation, convenience 3. Services (e.g., transport, shops)	School services to improve neighborhood
	Local community and charity (market)	1. Competitive demands 2. Potential enrollment 3. Voluntarily involve in school events 4. Funds by local charity	1. Fulfill demand of new courses/ curriculum 2. Fulfill demand of new service 3. Required knowledge to improve society (social prosperity, economic wealth, political stability) 4. Join in social responsibility activity
5. Immediate organizations after completion of education	Employers (companies)	Hiring students as employees in companies	Required knowledge (skills) for employment
	Higher educational institutions	Placement of students in institutions	Required knowledge (skills) to proceed to higher level
6. Other related parties	Student alumni	Student support (moral, advice)	Student engagement with alumni activities
	Scholarship sponsors	Student sponsorship	Students' engagement with foundation activities
	Suppliers	Provide services (e.g., food, transportation)	Commitment to utilize services
	Financial institutions	Financial assistance	Financial commitment

The purposes of society-oriented model for school education could be grasped from school and society point of views. Some of the purposes from society view are to gain benefits from the new value through recognition of new opportunities provided by school and to make use of existing value provided by schools through its student's skills or knowledge to improve the quality of living. Moreover, the model helps to create a self-belonging among society to willingly help schools to improve its operation and activities, and to recognize related parties to schools that can be worked together to improve the economic stability in local society. At the same time, the school may utilize the model to achieve uniform thinking within the school organization on who is the relevant and important stakeholders and to create a balance recognition to gain supports in realization of school objectives towards school excellence. In addition, school could find opportunities in creation of new value provided through school operation and output, and capture service satisfaction based on feedback as future improvements and new service offerings.

Appendix 2: Comparison of criteria and sub-criteria for the TQM model based on Deming criteria with general Deming Prize and Baldrige Excellence Framework (Education)

<u>1) Difference between the general Deming Prize framework and the proposed model</u>

Table II until Table XIII describe the criteria from the general Deming Prize framework and the proposed model, which include criteria and sub-criteria.

Table II: Descriptions of the A1 criterion and sub-criteria

Criteria from general Deming Prize
<u>Establishment of proactive customer-oriented business objectives and strategies</u> Under clear management belief, proactive customer-oriented business objectives and strategies has been established according to the management philosophy, industry, scale and environment, taking into account social responsibility of the organization. And the aspirations and future plans of the organization have been clearly spelt out.
Criteria based on TQM model based on Deming criteria for schools
<u>Establishment of proactive society-oriented school objectives and strategies</u> Formulation and establishment of society-oriented school objectives and strategies as school missions and visions; the school's future plans are clearly shared.
Sub-criteria
(i) School objectives are clearly formulated and well-established.
(ii) School strategies are clearly formulated and well-established.
(iii) Types of stakeholders involved in school are well-defined and understood within the school.
(iv) Business environment of school is well-understood during formulation of school objectives and strategies.
(v) School objectives and strategies are aligned with the management philosophy and business-environment condition.
(vi) School objectives and strategies are formulated based on a society-oriented.
(vii) School objectives and strategies are formulated considering the school's social responsibility.
(viii) Mid- and long-term of school objectives and strategies are clearly formulated and established.

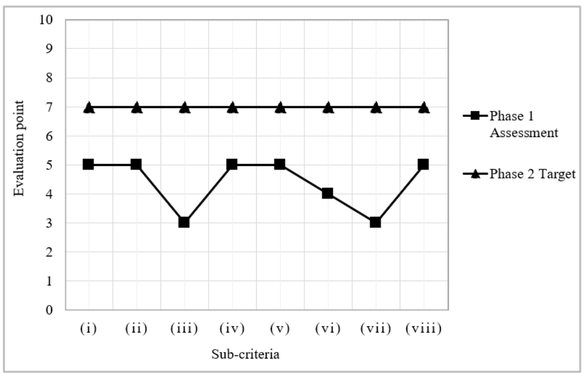


Figure I: Evaluation scores for the A1 sub-criteria

Table III: Descriptions of the A2 criterion and sub-criteria

Criteria from general Deming Prize
<u>Role of top management and its fulfillment</u> Top management is exhibiting leadership in formulation of proactive customer-oriented business objectives and strategies and implementation of TQM. It has insight concerning business objectives, strategies and environmental change and understands the importance of enhancement of organizational capabilities, human resource development and corporate social responsibility. It has understanding of and enthusiasm towards TQM.
Criteria based on TQM model based on Deming criteria for schools
<u>Role of school top management and its fulfilment</u> School top management, including the principal, are involved in formulating society-oriented school objectives and strategies, enhancement of school capabilities, and demonstration of enthusiasm for TQM implementation.
Sub-criteria
(i) Top management has insight concerning school objectives and strategies.
(ii) Top management aware of changes in educational business/service environment.
(iii) Top management shows leadership in formulation of society-oriented school objectives and strategies.
(iv) Top management demonstrates enthusiasm towards TQM.
(v) Top management reveals leadership in implementation of TQM.

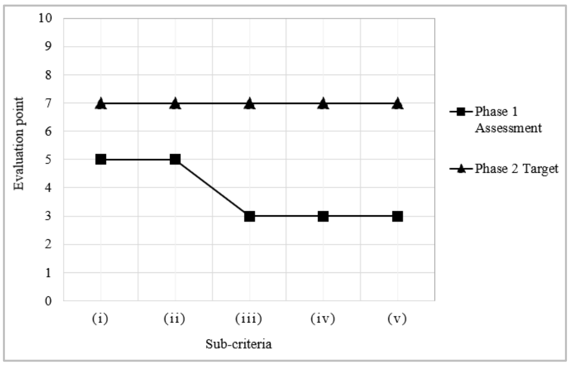


Figure II: Evaluation scores for the A2 sub-criteria

Table IV: Descriptions of the B1 criterion and sub-criteria

Criteria from general Deming Prize
<u>Organizational deployment of business objectives and strategies</u> Business objectives and strategies are being deployed throughout the organization and implemented in a united way based on total employee involvement, and close cooperation between departments and related organizations.
Criteria based on TQM model based on Deming criteria for schools
<u>School-wide deployment of school objectives and strategies</u> A systematic deployment of school objectives and strategies to achieve school excellence based on challenging targets. Deployment of objectives and strategies throughout the school organization based on cooperation among departments, faculties, and staff.

Table V: Descriptions of the B2 criterion and sub-criteria

Criteria from general Deming Prize
Establishment of proactive customer-oriented business objectives and strategies
Under clear management belief, proactive customer-oriented business objectives and strategies has been established according to the management philosophy, industry, scale and environment, taking into account social responsibility of the organization. And the aspirations and future plans of the organization have been clearly spelt out.
Criteria based on TQM model based on Deming criteria for schools
Creation of new educational values based on understanding of societal needs
Development and management of new educational services and school processes are proactively and effectively performed, aiming to create new educational values based on societal needs. Innovative technologies are utilized for the creation of new values in education.
Sub-criteria
(i) The needs of society-oriented parties are clearly identified.
(ii) Value, in terms of educational services or school processes, which has been provided by the school are clarified.
(iii) New value, in terms of educational services or school processes (e.g., curricular, course), is proposed based on societal needs.
(iv) A development process or workflow to introduce new value in terms of educational services or school processes is established and practiced.
(v) New value, in terms of educational services or school processes, is designed and developed based on the established development process or workflow.
(vi) Before its implementation, an internal review process at the school is carried out for detailed evaluation.
(vii) After its implementation, a review process to check the effectiveness of the new services or processes is carried out to get feedback from society.

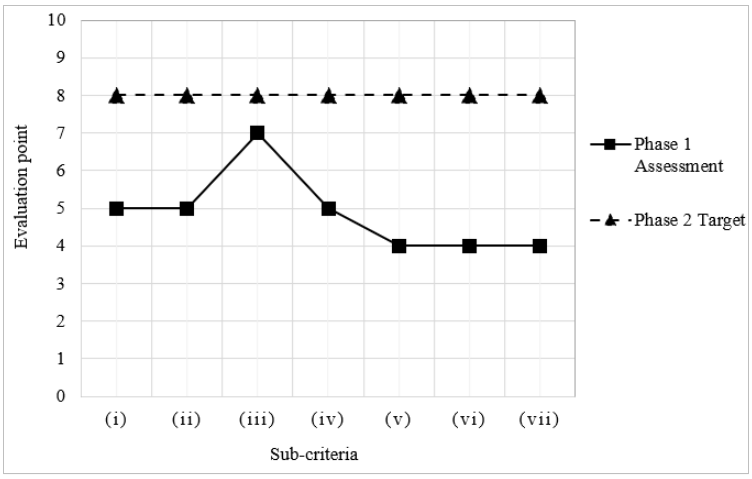


Figure III: Evaluation scores for the B2 sub-criteria

Table VI: Descriptions of the B3a criterion and sub-criteria

Criteria from general Deming Prize
Management and improvement of quality of products and services and/or work process Daily Management: There are few troubles in day-to-day operations through standardization and education & training and major operations in each department have been stabilized.
Criteria based on TQM model based on Deming criteria for schools
Management of quality in educational services and supporting work Implementation of subject-specific supporting work, educational management, and standardization of the school's routine work, especially in the core processes of education (teaching and learning). These activities are important for stabilizing the process of achieving the desired targets. This includes proper education and training for teachers and school staff in the provision of education to students. Implementation of subject-specific supporting work, educational management, and standardization of the school's routine work, especially in the core processes of education (teaching and learning). These activities are important for stabilizing the process of achieving the desired targets. This includes proper education and training for teachers and school staff in the provision of education to students.
Sub-criteria
(i) Missions and roles of educational management tasks in school processes are clarified and aligned with school objectives and mid-to-long-term planning. This includes the teaching and learning processes in classrooms to provide education to students.
(ii) Educational management tasks in school processes are analyzed according to their functions in task development until these reach a procedure level.
(iii) Educational management tasks are clarified through a process-based (input-process-output) approach with an established workflow.
(iv) Standardization of educational management tasks are performed through determination of important factors to support these processes, the creation of standards, and training.
(v) A detection system for conditions that are abnormal (unusual conditions') in educational management tasks is realized through visualization and frequent monitoring.
(vi) Educational management tasks are carried out according to agreed-upon standards, with continuous monitoring by the school.
(vii) Detection of unstable conditions in educational management tasks is shared and remedial action is immediately conducted.
(viii) Investigation is carried out to determine the root cause of unstable conditions, including prevention of their recurrence.
(ix) Management of quality in work processes is ingrained in the school to foster the development of a quality culture.

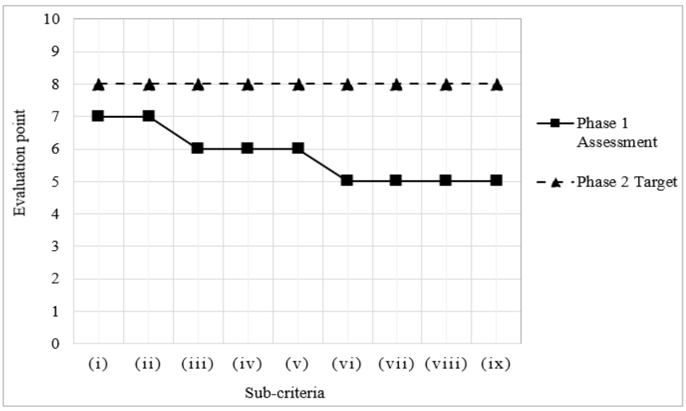


Figure IV: Evaluation scores for the B3a sub-criteria

Table VII: Descriptions of the B3b criterion and sub-criteria

Criteria from general Deming Prize
Management and improvement of quality of products and services and/or work process Continuous improvement: Improvements in quality of products and services and/or work processes are being carried out in a planned and continual manner and claims and defects in the market and/or next-processes are decreasing or are being maintained at an extremely low level. Customer satisfaction level has improved or is being maintained at an extremely high level.
Criteria based on TQM model based on Deming criteria for schools
Continuous improvement in quality of all school services and processes Implementation of continuous improvement in all school services and processes to systematically improve the quality of the school. Through continuous improvement activities, problematic areas in the provision of services, length of execution of routine tasks, and complaints from external parties may be reduced. This may increase the level of service satisfaction and happiness resulting from all school processes.
Sub-criteria
(i) Through PDCA cycle, process or service that requires improvement is identified and analyzed through appropriate sources (e.g., data from actual workplace, society-oriented parties' complaints)
(ii) By managing the previous process and understanding current situation during which data are gathered, analysis of root cause is carried out using quality control tools.
(iii) Improvement ideas for the problem are planned in a comprehensive way.
(iv) Improvement plans for the problem is proactively executed.
(v) Implementation of improvement idea is monitored closely to check if it has brought about the desired improvements.
(vi) The effective new method is introduced and practiced for sustained improvement.
(vii) All level of leaders and employees in school involve in continuous improvement activities.
(viii) Through continuous improvement process, the complaints towards school have reduced at an extremely low level.
(ix) Through continuous improvement process, the customer satisfaction towards school has improved at an extremely high level.

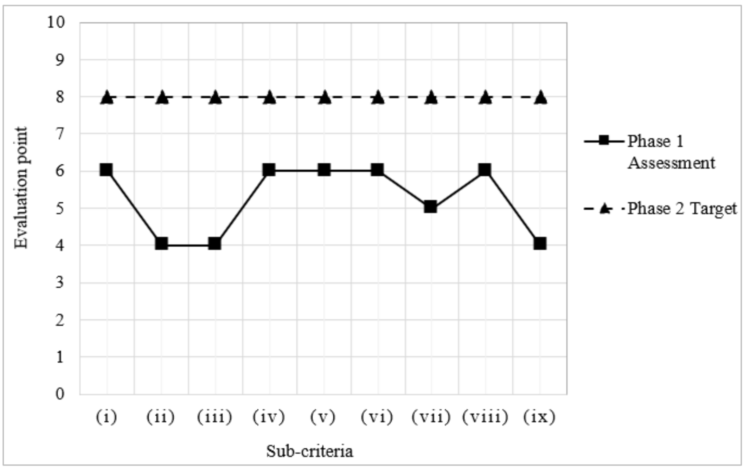


Figure V: Evaluation scores for the B3b sub-criteria

Table VIII: Descriptions of the B4 criterion and sub-criteria

Criteria from general Deming Prize
Establishment and operation of cross-functional management systems such as quality, quantity, delivery, cost, safety, environment, etc. across the supply chain Cross-functional management systems necessary for the organization are being established and operated suitably across the supply chain headed to customers including partners and related organizations and are effective in achieving the objectives in the changing business environment in a rapid and reliable manner.
Criteria based on TQM model based on Deming criteria for schools
Establishment of school-wide quality assurance system Establishment of a school-wide quality assurance system to improve overall education quality, especially in the provision of values in learning. For teaching staff, this covers the process of student admissions, placement in classrooms, and performance monitoring and graduation. Various departments must be involved to assure the highest quality of learning for students. The management system for operations related to safety, cost, environment, and school suppliers should also be involved
Sub-criteria
(i) A quality assurance system (cross-functional management) is available and operated fits the scale of the school.
(ii) A quality assurance system works horizontally across department/faculty and functions.
(iii) A quality assurance system involves employees with relevant skills and knowledge available in different functions (e.g., admission, academic, examination, quality assurance)
(iv) A quality assurance system includes the supply chain or/and business partners of school (e.g., transportation, safety environment, book supplies, food supplies).
(v) A quality assurance system operates effectively and fast in problem solving for a smooth school operation.
(vi) Quality assurance activities support to improve and assure overall educational quality.

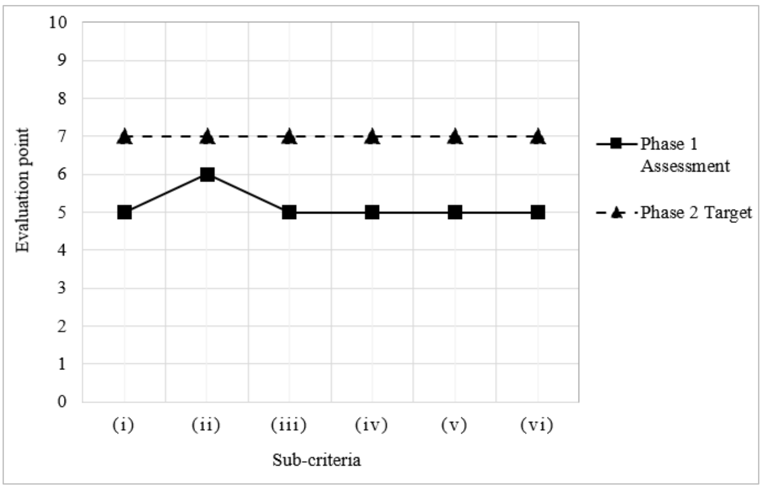


Figure VI: Evaluation scores for the B4 sub-criteria

Table IX: Descriptions of the B5 criterion and sub-criteria

Criteria from general Deming Prize
Collection and analysis of information and accumulation and utilization of knowledge Collection and analysis of Information from the market and within the organization and accumulation and use of knowledge necessary for operations is being carried out in an organized manner. In addition, such information is useful in creation of new values, management and improvement of products, services and/or operational quality and establishment and operation of cross-functional management systems.
Criteria based on TQM model based on Deming criteria for schools
Collection and analysis of information and utilization of accumulated knowledge Establishment of a system to systematically collect and analyze feedback from members of society. This also extends to the voices of parents, students, and school staff, and includes analysis of student performance and utilization of information technology (IT). Such information and accumulated knowledge are employed to create new values in educational services and operations to improve learning quality
Sub-criteria
(i) Information from external society-oriented parties (e.g., parents, local community) is collected and analyzed through an appropriate management system.
(ii) Information within school organization (e.g., teachers, employees) is collected and analyzed through an appropriate management system.
(iii) Knowledge is managed in a systematic way to support the underlying school objectives and strategies.
(iv) Accumulated knowledge and information are integrated with external society-oriented parties' feedback to create new values (e.g., new curriculum offers).
(v) Accumulated knowledge and information are integrated with feedback within school organization to create new values (e.g., enhancement of work process).
(vi) The information and knowledge management are utilized to benefit the operation of a quality assurance management system.

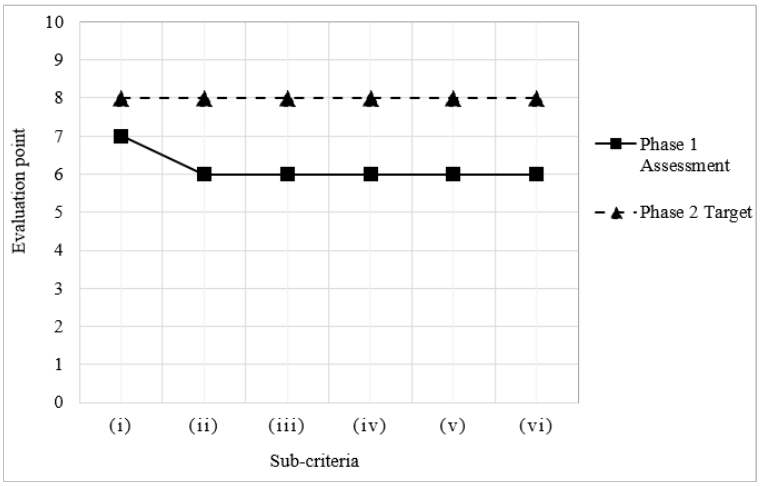


Figure VII: Evaluation scores for the B5 sub-criteria

Table X: Descriptions of the B6 criterion and sub-criteria

Criteria from general Deming Prize
Development and active utilization of human resource and organizational capability Development of human resource and organizational capabilities is being carried out in a planned manner and it is useful in realization of business objectives and strategies and implementation of TQM, and activation of people and organization that supports them.
Criteria based on TQM model based on Deming criteria for schools
Development of school staff and proactive utilization of school capabilities Development of skills and relevant abilities in teachers and school staff in supporting departments in order to provide better education and embed a culture of quality at the school. Furthermore, school capabilities are used to realize school objectives and strategies in the TQM implementation
Sub-criteria
(i) A recruitment, selection and placement system in relation to building up a quality institution are practiced and established.
(ii) Training and development for skills enhancement in job performing, including quality-related competences training, are well-executed.
(iii) A compensation/incentives (rewards and recognition) system as motivation to employees is established for enhancement of employees' motivation.
(iv) An effective performance appraisal or performance evaluation system for evaluation of all employees is established for enhancement of employees' motivation.
(v) A human resource development is practiced through alignment of employees' development with organization's goal.
(vi) The organizational capability is utilized towards better education through TQM implementation.

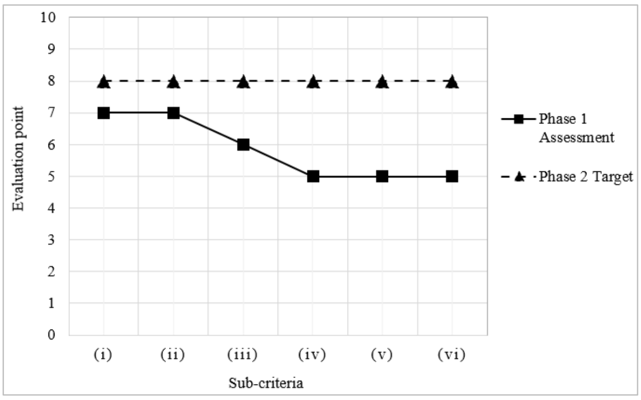


Figure VIII: Evaluation scores for the B6 sub-criteria

Table XI: Descriptions of the B7 criterion and sub-criteria

Criteria from general Deming Prize
Initiatives for social responsibility of the organization The organization is aware of its role and responsibilities as a member of the society and has established specific indicators in this regard and is adopting initiatives proactively (for instance, environmental preservation, regional contribution, fair operating practices, respect for human rights, information security, etc.) according to its management philosophy, type of industry, business scale and business environment.

Criteria based on TQM model based on Deming criteria for schools
Initiating and fulfilling the school's social responsibility Being active in performing school roles and obligations as part of society through social responsibility activities that contribute to the well-being of the immediate community. The school plays a role in advocating environmental sustainability and promoting ethical operating practices such as transparency and integrity.
Sub-criteria
(i) School understands their roles and obligation as a member of society in contributing to the well-being of their immediate community.
(ii) School plays its role to advocate of environmental sustainability (e.g., environmental preservation).
(iii) School is promoting ethical business practices (e.g., fair operating practices, transparency, integrity, equality).

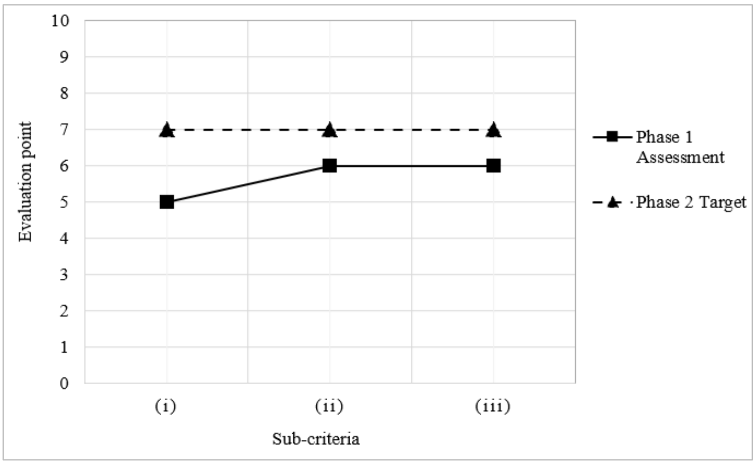


Figure IX: Evaluation scores for the B7 sub-criteria

Table XII: Descriptions of the C1 criterion and sub-criteria

Criteria from general Deming Prize
Effects obtained regarding business objectives and strategies through utilization and implementation of TQM The organization has obtained effects on business objectives and strategies through suitable utilization and implementation of TQM.
Criteria based on TQM model based on Deming criteria for schools
Realization of the effects of objectives and strategies through TQM implementation The school has realized the effects of its society-oriented school objectives and strategies through the implementation and utilization of TQM suitable to the school context.
Sub-criteria
(i) School has established systematic management systems and a total management system suit to business-environment of school.
(ii) Positive results are obtained from a suitable utilization and implementation of TQM (e.g., teamwork, continuous improvement activities, employees' empowerment, ingrain of quality culture).

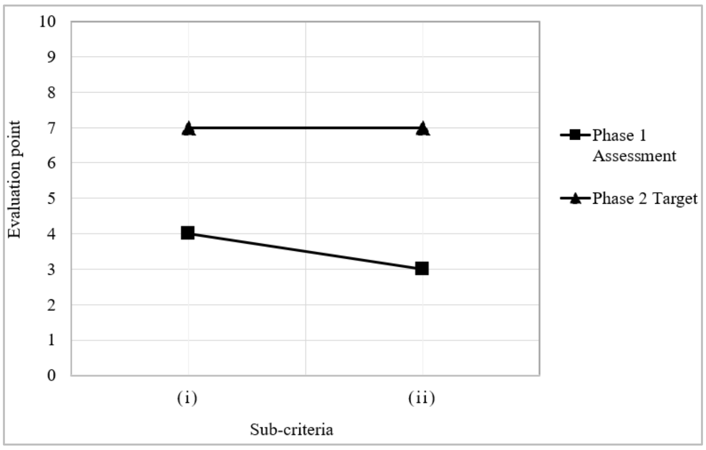


Figure X: Evaluation scores for the C1 sub-criteria

Table XIII: Descriptions of the C2 criterion and sub-criteria

Criteria from general Deming Prize
Outstanding TQM activities and acquisition of organizational capabilities The organization has obtained effects in the core areas for the realization of business objectives and strategies based on outstanding TQM activities regarding content and/or application of TQM and has acquired organizational capabilities necessary for its future sustainable growth.
Criteria based on TQM model based on Deming criteria for schools
Exceptional TQM activities and acquisition of school capabilities Through exceptional TQM implementation, the school has had positive effects on its core processes of teaching and learning to realize school objectives and strategies and to acquire capabilities necessary for future growth.
Sub-criteria
(i) School has obtained effects in core process (e.g., internal operating process, customer satisfaction, employee satisfaction, financial performance) based on outstanding TQM activities.
(ii) Through TQM implementation, the school has acquired the organizational capabilities in term of tangible/ intangible resources (e.g., facilities, financial resources, collective knowledge, high-skill employee)

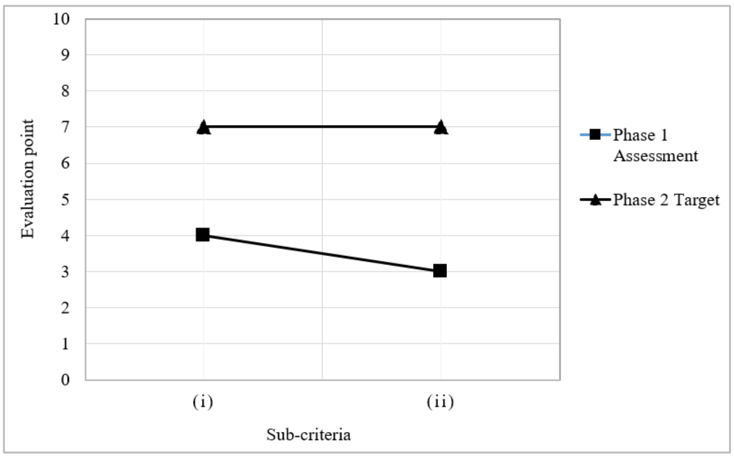


Figure XI: Evaluation scores for the C2 sub-criteria

2) Detailed descriptions of Baldrige Excellence Framework (Education)

Based on NIST (2015), by challenging to answer the questions that make up the Education Criteria for Performance Excellence, the organization explores how they accomplish what is important to them. The Baldrige Excellence Framework is used for improvements, whether the organizations just want to learn about the framework or if the organization is ready to assess the organization using Baldrige, or if the organization is ready for the external assessment. The questions represent seven critical aspects of managing and performing as an organization: 1. Leadership, 2. Strategy, 3. Customers, 4. Measurement, analysis, and knowledge management, 5. Workforce, 6. Operations, and 7. Results. An example is selected for detail description, which is Strategy. The Strategy category asks how the organization develops strategic objectives and action plans, implements them, changes them if circumstances require, and measures progress.

Criteria 2: Strategy

2.1 Strategy Development: How do you develop your strategy?

a. Strategy Development Process

(1) Strategic Planning Process

How do you conduct your strategic planning? What are the key process steps? Who are the key participants? What are your short- and longer-term planning horizons? How are they addressed in the planning process? How does your strategic planning process address the potential need for

- transformational change and prioritization of change initiatives
- organizational agility, and
- operational flexibility?

(2) Innovation

How does your strategy development process stimulate and incorporate innovation? How do you identify strategic opportunities? How do you decide which strategic opportunities are intelligent risks for pursuing? What are your key strategic opportunities?

(3) Strategy Considerations

How do you collect and analyze relevant data and develop information for your strategic planning process? In this collection and analysis, how do you include these key elements?

- Your strategic challenges and strategic advantages

- Risks to your organization's future success
- Potential changes in your regulatory environment
- Potential blind spots in your strategic planning process and information
- Your ability to execute the strategic plan.

(4) Work Systems and Core Competencies

What are your key work systems? How do you make work system decisions that facilitate the accomplishment of your strategic objectives? How do you decide which key processes will be accomplished by external suppliers and partners? How do those decisions consider your core competencies and the core competencies of potential suppliers and partners? How do you determine future organizational core competencies and work systems?

b. Strategic Objectives

(1) Key Strategic Objectives

What are your organization's key strategic objectives and timetable for achieving them? What are your most important goals for these strategic objectives? What key changes, if any, are planned in your educational programs and services, customers and markets, suppliers and partners, and operations?

(2) Strategic Objective Considerations

How do your strategic objectives achieve appropriate balance among varying and potentially competing organizational needs? How do your strategic objectives

- address your strategic challenges and leverage your core competencies, strategic advantages, and strategic opportunities;
- balance short- and longer-term planning horizons; and
- consider and balance the needs of all key stakeholders?

2.2 *Strategy Implementation: How do you implement your strategy?*

a. Action Plan Development and Deployment

(1) Action Plans

What are your key short- and longer-term action plans? What is their relationship to your strategic objectives? HOW do you develop your action plans?

(2) Action Plan Implementation

How do you deploy your action plans? How do you deploy your action plans to your workforce and to key suppliers, partners, and collaborators, as appropriate, to ensure that you achieve your key strategic objectives? How do you ensure that you can sustain the key outcomes of your action plans?

(3) Resource Allocation

How do you ensure that financial and other resources are available to support the achievement of your action plans while you meet current obligations? How do you allocate these resources to support the plans? How do you manage the risks associated with the plans to ensure your financial viability?

(4) Workforce Plans

What are your key workforce plans to support your short- and longer-term strategic objectives and action plans? How do the plans address potential impacts on your workforce members and any potential changes in workforce capability and capacity needs?

(5) Performance Measures

What key performance measures or indicators do you use to track the achievement and effectiveness of your action plans? How does your overall action plan measurement system reinforce organizational alignment?

(6) Performance Projections

For these key performance measures or indicators, what are your performance projections for your short- and longer-term planning horizons? How does your projected performance on these measures or indicators compare with your projections of the performance of your competitors or comparable organizations and with key benchmarks, as appropriate? If there are gaps in performance against your competitors or comparable organizations, how do you address them?

b. Action Plan Modification

How do you establish and implement modified action plans if circumstances require a shift in plans and rapid execution of new plans?

3) Comparison between the general Deming Prize, Baldrige Excellence Framework (Education) and the TQM model based on Deming criteria for school education

Baldrige Excellence Framework (Education) questions typically begin with ‘how’ and ‘what’ to assist applicants in explaining their processes in relation to the criterion for ‘Strategy’. For example, based on the subsequent questions that follow the main question, ‘how do you conduct your strategic planning?’, the organization or applicant must explain how they implement strategic planning and give further information on key processes, key participants, and so forth. The Baldrige examiners can assess the applicant using the detailed and prescriptive questions when they seek for an external evaluation. Accordingly, this framework tends to encourage gradual improvements and the applicant assesses their organization based on existing performance and results.

On the other hands, JUSE (2021) asserts that the Deming Prize examination does not require that applicants adhere to a quality management framework offered by the Deming Prize Committee. Instead, the applicants are expected to understand their current condition, set their own themes and goals, and transform the organization. Not only the results achieved and the processes used, but also the effectiveness expected in the future is subjects for the examination. The general Deming framework places emphasis on to-be-status. For example, in B1 criteria of ‘Organizational deployment of business objectives and strategies’, the aim to be achieved by the applicant is that the ‘business objectives and strategies are being deployed throughout the organization and implemented in a united way based on total employee involvement, and close cooperation between departments and related organizations’. Once the applicant has achieved this, there are a variety of detailed questions that can be asked, such as ‘how the applicant feels that this approach is appropriate for achieving the organization's objectives’, and so forth. Although it can be difficult for an organization to see their own deficiencies and come up with remedies, it helps the applicant in developing organizational skills and ability for the present and the future. This explains why the Baldrige framework is more prescriptive than the Deming framework.

The TQM model based on Deming criteria is based on the main Deming framework's philosophy and includes extensions of sub-criteria for each criterion to guide schools on how to conduct TQM activities in a systematic way. For example, in B1 criterion of ‘School-wide deployment of school objectives and strategies’, the PDCA cycle, from which the sub-criteria are generated, is a fundamental management cycle that schools must adopt in order to be prepared to handle any challenging future objectives. For instance, the sub-criteria (iii) in B1 of "School-wide deployment of school objectives and strategies" emphasizes the breakdown of the primary objective into smaller objectives, and these objectives must be deployed with the proper methods and a distinct target. This will allow schools to evaluate their own practices and determine whether they have clear methods or means to carry out activities to accomplish the smaller target, by using one specific case. By doing this, schools are compelled to find their own themes for implementing a challenging policy while also getting additional guidance on how to implement the policy in a systematic way.

Appendix 3: Questions for interviews and surveys

i) Questions for semi-structured interviews with school management

Interview question 1 (Iq1):

The TQM model based on Deming criteria is expected to provide hints to schools to enhance organizational ability in education management, not only for the present but also for the future, towards achieving school excellence. Based on TQM implementation in your school, have you experienced this benefit?

Interview question 2 (Iq2):

For society-oriented consideration of school objectives and strategies, the expected benefit of the model is to provide guidance to schools about appropriately recognizing stakeholders. In your experience,

i. Did it help you to recognize the important stakeholders in your school?

ii. Do you think that your school policy or school direction has appropriately considered such stakeholders?

iii. Do you think that the model gives hints on how to create a balanced stakeholder recognition that helps you gain support from stakeholders to realize your school objectives?

Interview question 3 (Iq3):

The TQM model based on Deming criteria is also expected to provide ideas to schools to link the results and effects of TQM implementation with the society-oriented school objectives and strategies. Did you experience this benefit?

Interview question 4 (Iq4):

The expected benefits of these sub-criteria are to guide schools on carrying out systematic processes to implement school objectives and strategies based on sub-criteria (i) to (vi) for B1.

Based on the activities and evaluation process, what is your opinion or impression of the model? Did you gain any benefits from these activities?

ii) Survey questions to school's middle management and staff

i. Closed-ended questions:

Response scale: 1 – Strongly disagree, 2 – Agree, 3 – neither agree nor disagree, 4 – agree, 5 – strongly agree)

Survey question 1 (Sq1): The model is helpful in offering ideas and useful suggestions on how to carry out TQM at your school.

Survey question 2 (Sq2): The model provides visibility on school strengths and weaknesses for future improvements.

ii. Open-ended question: What is your opinion of the TQM Deming model for schools?

Appendix 4: A daily management model for teaching and learning in school education

i) Generation of the new model for daily management in teaching and learning process

A model for teaching and learning in classrooms is designed for the usage in the core process of education. It is modified from the JSQC Standard of Guidelines for Daily Management (JSQC, 2014) and made congruent into the educational context, particularly in the classrooms. This model is a combination of two criteria of the general model; the fundamental of daily management, which includes the standardize-do-check-act (SDCA) cycle with its single element definition, and the elements under a general way to implement the daily management in a single job unit. When the elements are combined, a single diagram with a description table is designed, and those overlapped elements that appeared in the model is positioned at the appropriate step for better comprehension. The next step is to identify element for every SDCA step, which is relevant to education, and to prioritize the elements based on educational process. The translation of the general process into educational terminologies and understanding to suit to teaching and learning in a classroom setting is the most critical step in the creation of this model. Some elements are irrelevant to the education process, and not indicated in this model.

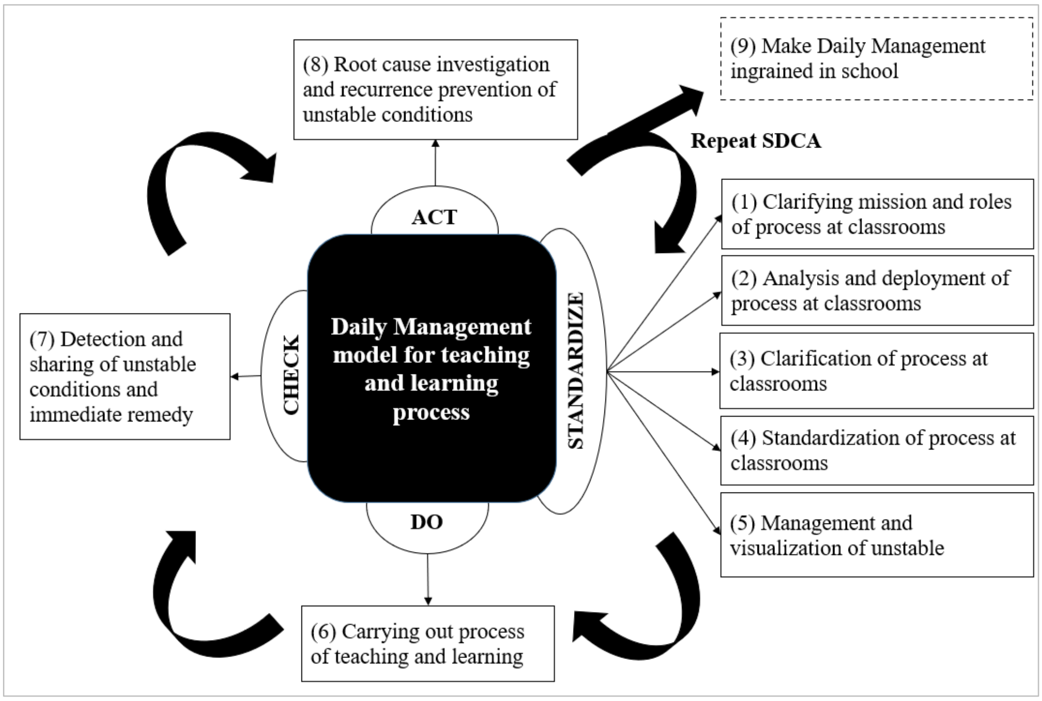


Figure XII: A model of daily management for teaching and learning at classrooms

The model is created as a diagram with a table to explain the contents further. In Figure XII, SDCA is called as 'step', and (1) to (9) is called as 'Element'. In this model, 'clause' and 'sub-clause' are modified into 'element' and 'sub-element' to show the components' characteristics of the activities. The total of nine elements and twenty-seven sub-elements are further described in Table XIV, which 1.1 to 9.2 are called as 'sub-element'. The SDCA cycle will keep repeating until it is established. For this, element (9) is vital to ensure that Element (1) to (8) are well maintained; thus, the efforts for system reviewing, human development, and creating cultures are important.

Overall, the purpose of the model is to enhance process effectiveness in education processes through standardization activities based on the SDCA cycle. This model helps to stabilize the processes in producing output and the output itself through recognition, visualization, and management of important related factors and activities. Besides, it is to improve effectiveness in teaching and learning processes in classrooms based on lesson delivery, students' assessment, and target output, including the abnormality recognition, through the management of standards. As a long-term goal, through the implementation of this model, it helps to ingrain the quality culture among school staff in specific and entire school processes. The model scope is mainly for the usage in the teaching and learning that happens in the classroom and focuses on teacher and school administrators as the process owners.

Table XIV: A description table of daily management model for teaching and learning process at classrooms

Element	Sub-element
(1) Clarifying mission and roles of process at classrooms	1.1 Clarification of mission and role of a departmental task based on current responsibility and aligned to school mission and mid-to-long-term plan. (**)
	1.2 Identifying each curriculum/course required to be managed. (**)
(2) Analysis and deployment of process at classrooms	2.1. Definition of task in classroom by breaking down process based on functional expression (e.g., carry out lesson based on lesson plan). (**)
	2.2 Analyzing task in classroom according to its function for task deployment until it reaches procedure level (e.g., Level 1: teaching English Cambridge for Grade 9 -> Level 2: teaching Chapter 1 of English Cambridge for Grade 9; Level 3: preparing lesson plan, assessing student understanding).(**)

(3) Clarification of process at classrooms	3.1 Clarification of teaching and learning process flow in relations to result in 2.2 by using process chart symbols and arrows. (e.g., process flow of teaching Chapter 1 of English Cambridge for Grade 9).
	3.2 Clarification of input and output of individual process by recognizing important factors at each step.
	3.3 Clarification of process ownership among teachers, school administrator and students who involve directly in teaching and learning.
(4) Standardization of process at classrooms	4.1 Defining of important factors to support teaching and learning and classifying them according to its categories (e.g., method: monitoring, feedback; manpower: teacher, student; material: lesson plan).
	4.2 Developing the standards used in classroom to confirm the output can be produced whenever work by using this standard.
	4.3 Providing education and training to employees to provide skills and competencies in carrying out tasks.
	4.4 School leaders ensure relevant employees of the process to develop relevant documents as standards, including its revision. (**)
	4.5 Standards are revised from time to time to make it more effective in carrying out teaching and learning process and must be carefully recorded. (**)
	4.6 Clarifying relevant changes according to important factors determined in 4.1, called as change management. (**)
(5) Management and visualization of abnormality / unstable condition	5.1 Determination of control item as a rating scale to manage the achievement of the objectives in teaching and learning teaching and learning process.
	5.2 Determination of control level according to control item of 5.1.
	5.3 Determination of frequency of data collection to check the control item and control level. (*)
	5.4 Visualization of unstable condition. (*)
(6) Carrying out process of teaching and learning teaching and learning at classrooms	6.1 Carrying out process according to agreed standards.
	6.2 Monitoring the achievement result for the process.
	6.3 Planning observation whether standards is followed or not, and results achieve target or not.
(7) Detection and sharing of abnormality/ unstable condition and immediate remedy	7.1 Occurrence of unstable condition is detected and checked.
	7.2 Detected unstable condition is shared and freely communicated within school organizations. (**)
	7.3 Immediate remedial action for unstable condition. (*)

(8) Root cause investigation and recurrence prevention for unstable condition	8.1 Investigating root cause of unstable condition by using QC tools and conduct improvements.
(9) Make daily management ingrained in school	9.1 Preparing and reviewing systems and relevant supporting tools/ facilities to ensure it is effectively working. (*)
	9.2 Human resources development and building up culture. (*)
	9.3 School management responsible to ensure all parties are performing their task and what is happening at school place. (*)

Note: The priority is classified into three categories; 'important sub-element to achieve its purpose', 'Sub-element may give impact to achieve its purpose' which is marked as (*) with light grey row, and 'Sub-element may give impact to achieve its purpose, but less important' with (**) mark in the dark grey row.

ii) Prioritization of elements to achieve process stabilization in teaching and learning

In order to effectively use the daily management model in the teaching and learning process in classrooms, the prioritization needs to be set. The priority is essential to allow sufficient time and readiness for school management and administrators to implement it at a more consistent speed gradually. In this research, three categories of prioritization are introduced to ensure daily management activities are smoother and may give ample time for teachers and school administrators. In Table XIV, the priority is classified into; a) important sub-element to achieve its purpose, b) sub-element that may give impact to achieve its purpose, which is marked as (*), and c) sub-element that may give impact to achieve its purpose, but less important' with (**). Daily management requires a long time to implement, especially in the Standardize stage, where many of the creation of the standards takes a longer time. Although it seems time-consuming for this stage, as highlighted by Ohno (1995), workers should start by following standard operations that available now, and as they proceed, the expectations for improvements will arise from the aspects of the operation by incorporating new ideas. Based on this model, there are a total of twenty-seven sub-elements to be carried out by teachers and school administrators. Ideally, schools should implement all elements and sub-elements at once as it provides steps and sequences for each task. However, it is worried the efforts stop at midways as there are many activities to carry out at once.

For a) category of 'the important element to achieve its purpose', thirteen sub-elements are categorized under this prioritization category, and these are the minimum steps or activities that school may carry out in order to start with the process stabilization in teaching and learning in classrooms. Eight of the sub-elements are categorized under Standardize steps, which could be described as the preparation steps before carrying out the daily process. This priority category contains the definitions of the teaching and learning process and to see each chapter lesson in every classroom as one process, and further break it down into steps for one classroom lesson. In teaching and learning, the idea of having a process flow itself is not commonly practiced as a teacher refers to the written standard as a guideline and be more flexible to modify the lesson plan depending on student response to carry out lessons (Boyd, 2012).

For b) category of 'elements that may give impact to achieve its purpose', six sub-elements are categorized under this category. It is also essential to carry out these sub-elements after completing a) as it gives more value-added to those elements. It gives more structure to sub-elements in a) towards a process stabilization in managing quality in education. By implementing the sub-elements in b), everyone who involves in the process will experience the quality management in daily system, and gradually it becomes a quality culture in the school.

For c) category of 'elements that may give impact to achieve its purpose, but less important', eight sub-elements are categorized under this category, and by carrying out these activities, it will complete the whole activities in the daily management model for the teaching and learning process. The activities emphasize the comprehensive management of routine work in classrooms, for example, clarifying the job purpose to align with school objectives. Elements and sub-elements in this category are considered less priority based on the importance of the fundamental structure; however, it helps to connect the links between each element. Furthermore, less priority definition is in terms of a time frame to give school ample time to carry out more priority activities first. Several elements are chosen for a detailed description in section iii).

iii) Detail description of selected important elements in teaching and learning process stabilization

Taking examples of three sub-elements, 3.1 in element (3) of clarification of processes in classrooms, and 5.1 and 5.2 under the element of (5) Determination of specific target value and its range when a target is not met (abnormality). For the sub-element 3.1, the process flow is an essential element to understand the entire tasks in sequencing steps when carrying out a lesson in a classroom. This step is one of the fundamental processes to maintain operating conditions with important factors described in sub-elements 3.2 and 4.1; thus, the output is stable to meet the requirements. An example of a process flow in the teaching and learning process in a classroom is given below, and it could be a general process for any lesson in a classroom environment. From each step in this process, it can be seen that which step requires specific management, which later classified as important factors in supporting the core processes of education. This process flow will be fundamental to the next sub-element to further describe a process-based approach by clarifying input, work, and output of the teaching and learning process.

Among all five processes, 2), 3), and 5) seems necessary to be provided with further detailed procedures or flow of standards. For instance, in process 3), the teacher may need to ensure the lesson plan, teaching aid, pedagogy, and textbooks are systematically prepared and managed for a specific subject matter and chapter. On the other hand, a teaching and learning process is an approach of student-centered teaching and learning environment (Baeten *et al.*, 2013), it requires a standard for students, too, as it can be seen at process 5). In this step, students' understanding is assessed through various formative assessments prepared by the teacher for the specific chapter's lesson. For instance, in Grade 9 mathematics, a student is completing chapter 1 of 'number systems'. There are fifteen chapters in the syllabus of one academic year for Grade 9 students, such as number systems, polynomials, and coordinate geometry. In process 5), the formative assessment is carried out before the class ends. If this is considered as the first process for Grade 9 students to learn mathematics, the next process is a chapter 2 lesson of 'polynomials'. Students can continue with 'polynomials' after they complete the lesson for 'number systems', and comprehension assessment.

Step 1: Students' attendance is checked
Step 2: Previous lesson is reviewed
Step 3: Lesson for the day is carried out
Step 4: Lesson is concluded
Step 5: Students' understanding is assessed

For the sub-element 5.1 and 5.2, these are described together as it has a close relationship with each other for the detection of abnormality. The difference between non-conformance and abnormality has to be understood to carry out these activities. Non-conformance is synonymous with defect or rejection that usually used in the manufacturing sector, and in a school context, it could be described as a student failure in a test, unsatisfactory homework completion, or absentee when a student does not come to school. For abnormality, it is the condition where the output result is not within the controlled condition, or not normal. A school needs to understand the 'normal' definition in each process to understand further what abnormality is. In reality, although teachers and students have followed standards accordingly, failure of students in tests or examinations will occur, and it is considered normal.

To detect abnormality in the teaching and learning, an example can be seen from the students' test scores management. A series of tests performed in one classroom could be analyzed, and a graph plots the average test scores of the total students. If the students' average test scores decline more than a specified minimum limit known as a lower control limit (LCL), it indicates that there is an abnormality. If the average of total students' test scores decreases but is still above LCL, it is in the controlled condition, and the teacher can monitor further without putting much effort into this area. The decision to define LCL is a crucial task to be accomplished in the 'Standardize' step, and the teacher might put more effort into deciding on this element.

From these examples, the understanding of daily management for process stabilization has to be carried out according to the sequence and prioritization category to ease the understanding and gain acceptance by teachers and school administrators. Although process stabilization in teaching and learning is time-consuming and requires untiring efforts from school milieus, it provides a new definition of a process-based approach in classrooms. Once schools implement the proposed model, it could help to stabilize the processes in classrooms, and from there, the classroom outputs are also stabled.

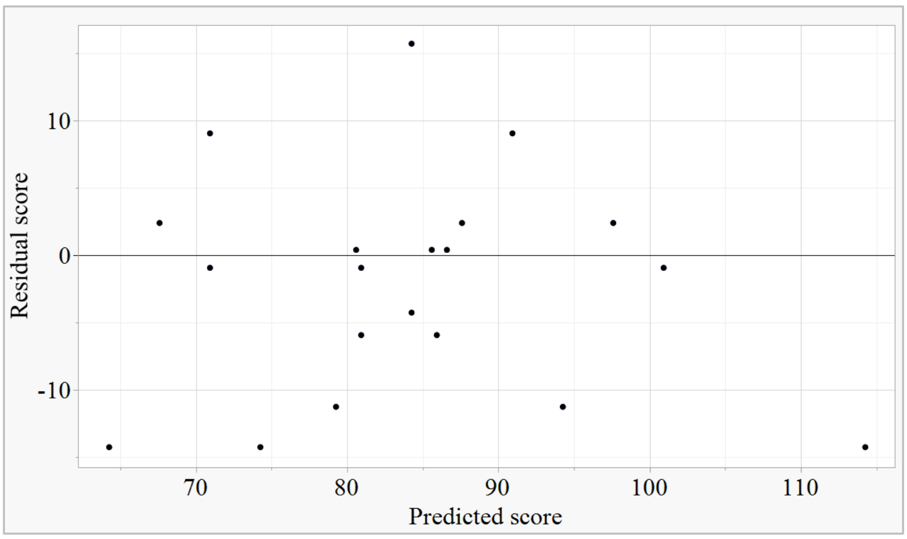
Appendix 5: Analysis of assumption of the model and possible reasons of high error occurrence

1) Validity of assumption of constant error variance for methods of detection of unusual and unexpected scores of individual students

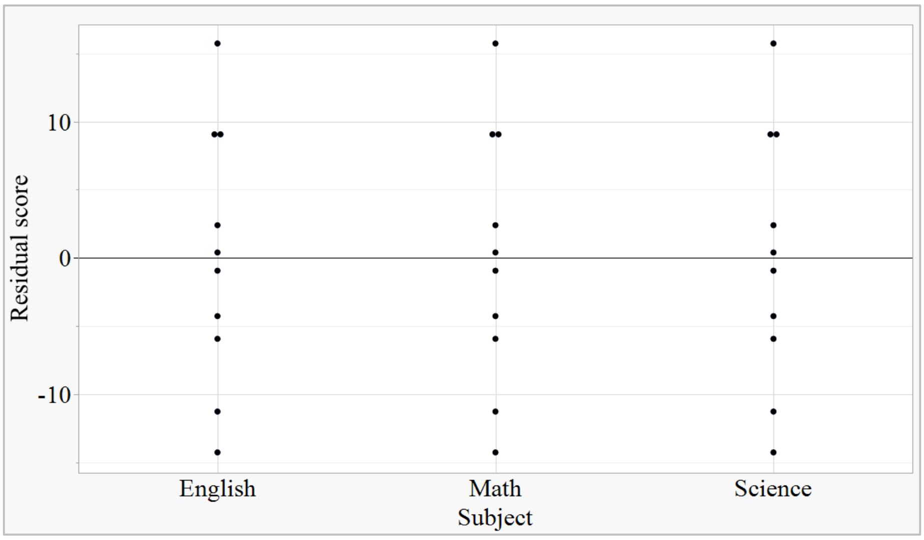
A residual, $Y_{ijt} - \hat{Y}_{ijt}$, is used to represent an error variance in Chapter 3 and 4. The violations of the basic assumptions and model adequacy can be investigated by the examination of residuals. A simple check is to plot the residuals versus the predicted values. This plot should not reveal any obvious pattern (e.g., Montgomery, 2012). If the graph showing a specific pattern such as an outward-opening funnel pattern or a double-bow pattern, it indicates that the variance of the errors is not constant (Montgomery *et al.*, 2012).

i) For dataset analysis in Chapter 3

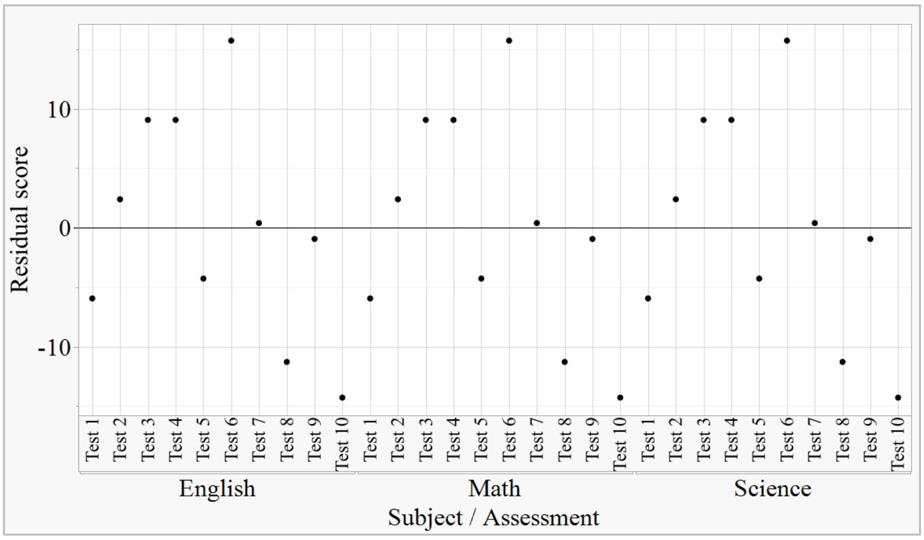
Figure XIII(a) demonstrates the residual score against the predicted scores regardless of students in 3 subjects and 10 assessments. The factor effect of Student is constant and the example is taken for student E dataset. Figure XIII(b) and Figure XIII(c) demonstrate each factor effect of Subject and Assessment×Subject at x-axis, respectively, for this dataset. The residual scores scattered between ± 15 points and focusing nearer to 0. Overall, there is no evidence that shows error variance is not constant.



(a)



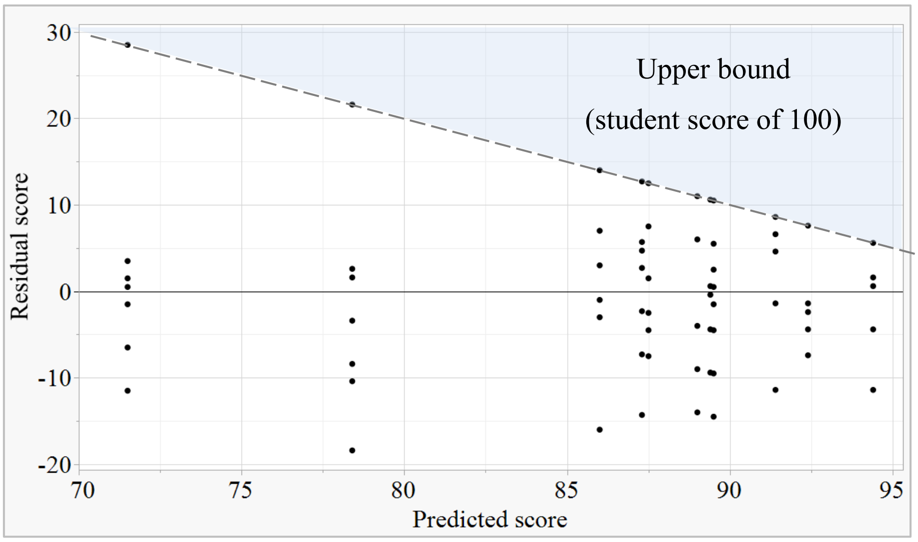
(b)



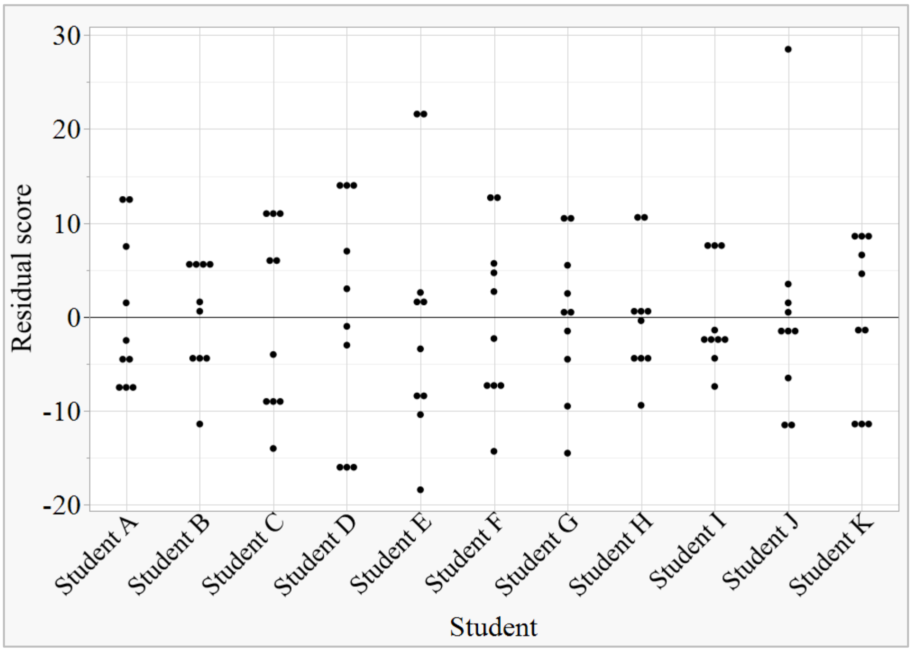
(c)

Figure XIII: Residual scores against (a) predicted scores for student E in mathematics, English and science for total of 10 assessments, (b) subject, and (c) assessment in subject.

Figure XIV(a) demonstrates the residual score against the predicted scores regardless of subjects for 11 students in 10 assessments. The factor effect of Subject is constant. Figure XIV(b) demonstrates a factor effect of Student at x-axis for the dataset in English. There is an area of upper bound that actual score is 100 when predicted score is other than 100. For example, based on previous performances, student F is predicted to score 87.3 in Assessment 4 and Assessment 5, and the student obtained 100 for both assessments. This is a nature of student scores that maximum score can only be achieved as 100 for the numerical grading system of 0 to 100. The residuals are scattered closer to 0 when predicted score is high. Also, there is a large residual when an actual score is 100 and predicted score is 71.5, considering the student scores in previous assessments. Although there is a sloping pattern due to the nature of actual student scores compared to its prediction, there is no evidence that shows error variance is not constant.



(a)



(b)

Figure XIV: Residual scores against predicted scores for English for all students in total of 10 assessments

Figure XV demonstrates the residual scores against the predicted scores regardless of assessments in a subject and individual student. Taking example of student C in science for 10 assessments, as the predicted score is fixed, which is equivalent to grand mean of 94.1, the residuals are plotted along 5 points when student C obtains 100. In this analysis, it shows no evidence that error variance is not constant.

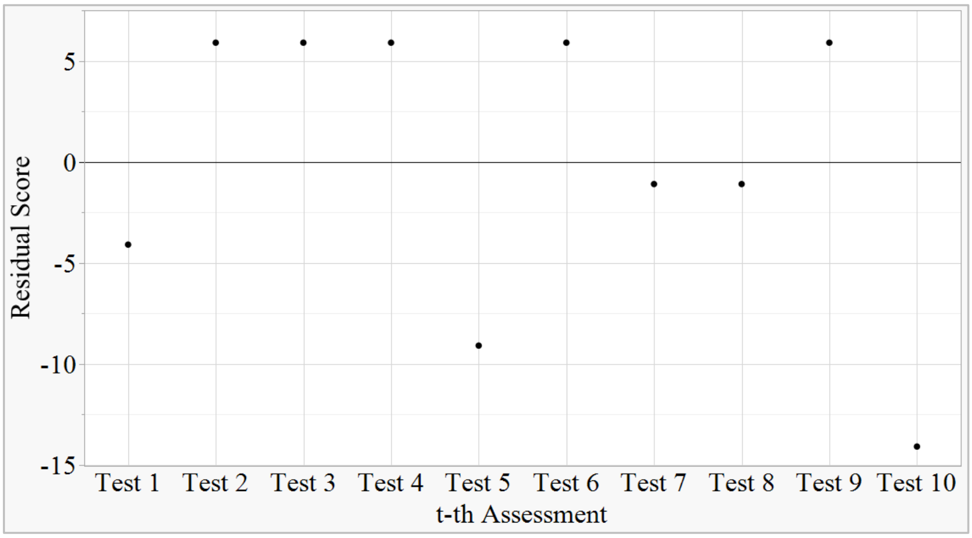


Figure XV: Residual scores against a factor effect of Assessment at x-axis for the dataset for student C for science in 10 assessments

In conclusion, based on scatter plots of residual scores and predicted scores, there is no evidence showing that error variance is not constant regardless of students, subjects, and assessments based on proposed model.

ii) For dataset analysis in Chapter 4

Figure XVI demonstrates the scatter plots of residual scores and predicted scores of past scores in semester 1 and semester 2 of 2020, current daily efforts in current semester of 2021, and trend of the current scores in current semester of 2021. Residuals are scattered more randomly within ± 10 when predicted score is low, and closer to 0 when predicted score is high. A residual is large for student M when the predicted score is 65.5, while the actual score is 50. Overall, there is no evidence that shows error variance is not constant.

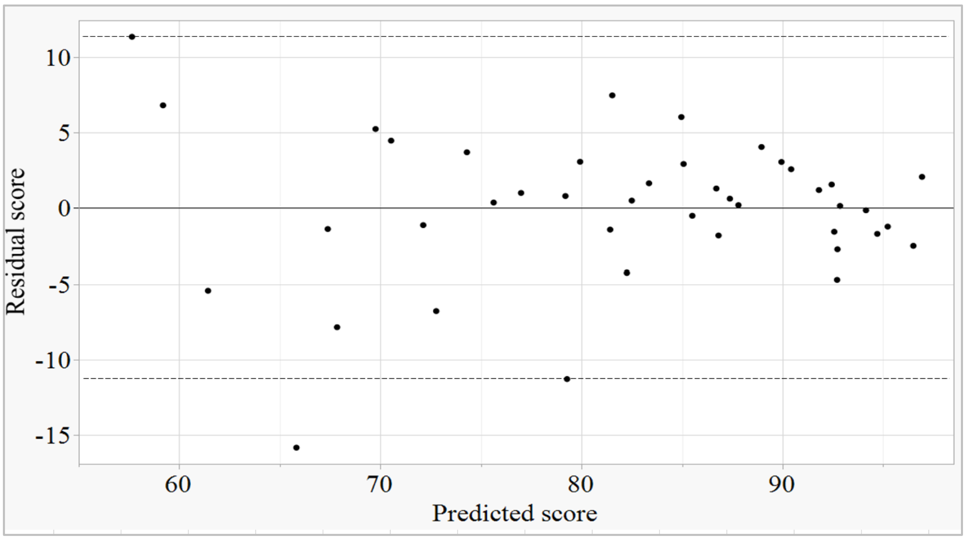


Figure XVI: Residual scores against predicted scores for past scores in 2020, current daily efforts in 2021 and trend of the current scores in 2021

2) Detailed analysis of potential reasons of high error occurrence

The detailed analysis of phenomenon occurred in Figure 4.3 is stated in this section. There are three possibilities of reasons that causing this occurrence.

i) Existence of influential outlier

Based on studentized residual analysis, unexpected (low) score detected for student M is an outlier. Although it is outlier, it may or may not influential to model. Generally, for Cook's distance values above 1.0, the data point is quite likely to be influential. In this analysis, Cook`s distance for the outlier is 0.74; thus, observation is not quite likely influential to the model. Besides, when removing outlier, $\hat{\sigma}$ is 7.68, as it is higher than 6.67. Considering some other factors, such as smaller number of data causing type II is likely to occur, it can be said that outlier is not quite influential to model.

ii) Magnitude of error differed based on subject

For subjects, based on unequal variance confirmation, such as Levene test, p-value is 0.59. Therefore, it is failed to reject the H_0 and there is insufficient evidence to claim that the variance between subjects is unequal. It is supported by the assumption of error variance is constant for all subjects has been checked based on scattered plot of residual against predicted score, stated at Figure XVI. Overall, there is no evidence that shows error variance is not constant in all subjects.

iii) Distribution of residuals scores that by chance are near to control limit

The phenomenon could be possible happen due to other than student M score in mathematics (outlier), there are residuals scores that by chance are near to UCL and LCL. Refer Figure XVII for the residual scores of each student in subjects. It shows that in subject such as science, the residual score for student L is -11.35 and for student M is 11.35, which are by chance near to UCL and LCL at $\pm 2\hat{\sigma}$. These scores are detected as normal scores. Due to these residuals, the error is slightly high despite of the predicted score and actual score are highly matched.

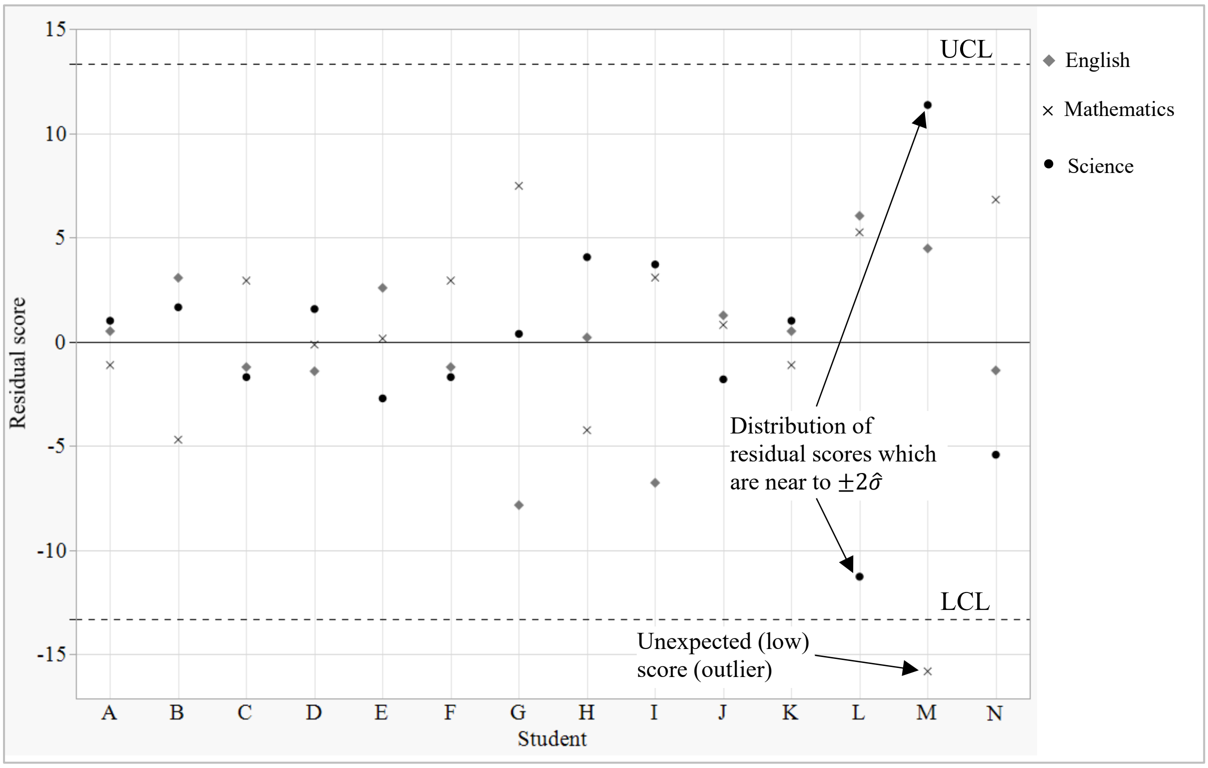


Figure XVII: Residual scores for each student in English, mathematics, and science

3) Validity of assumption of residuals are normally distributed for methods of detection of unexpected scores of individual students

Through the Q-Q plot, the distributions of the residuals follow a straight line. Therefore, it can be concluded that the residual is approximately normally distributed, see Figure XVIII.

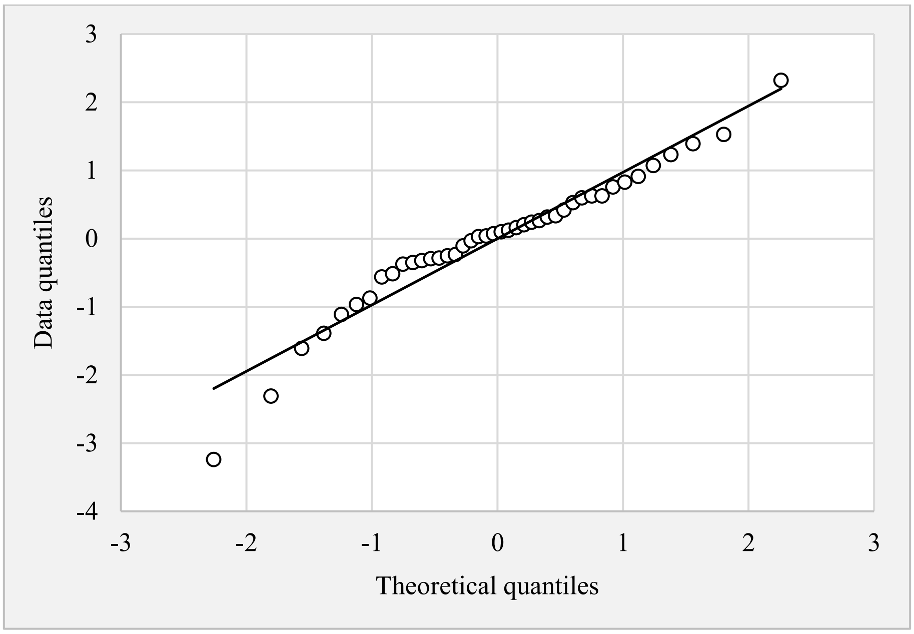


Figure XVIII: Residual analysis by Q-Q plot

Appendix 6: Analysis of dataset based on time series models

A time series is a set of observations $Y_1,\ Y_2, \dots, Y_t$ that are observed across an equally spaced time intervals. Time series errors have an auto correlated structure, which means they are correlated with themselves at different time intervals. Past data are used to predict and forecast future data in time series analysis. Time series analysis is used to study non-stationary data, or things that vary over time or are influenced by time. Time series analysis is commonly used in sectors such as banking, retail, and economics because currency and sales are always fluctuating.

In this analysis, several time series models based on autoregressive (AR), autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) are selected to analyze the prediction scores of t-th assessment. AR model is when a value from a time series is regressed on past values from that same time series. Autoregressive model of lag order p, often shortened to AR (p). ARMA is a mixture of autoregressive and moving average process, and it is the basic model for analyzing a stationary time series. Moving average is the dependency between an observed value and a residual error from a moving average model applied to previous observations. ARMA of order p, q is written as ARMA (p, q), which means an autoregressive model of order p is combined with a moving average model of order q. ARIMA is a model that includes autoregressive, integrated and moving average process. ARIMA utilizes differenced data to make the data stationary, which means there is a consistency of the data over time. This function removes the effect of trends or seasonality. The ARIMA model is denoted with the parameters (p, d, q), which parameter d is known as the degree of differencing and indicates the number of times the lagged indicators have been subtracted to make the data stationary. Detailed explanations of time series analysis are provided in standard textbooks, such as Montgomery *et al.* (2015).

i) For dataset analysis in Chapter 3

Daily assessment data from Assessment 1 to 9 are used as past data to predict scores of Assessment 10. For this analysis, it is assumed that the errors of student scores have auto correlated structure. Based on each individual student score in a subject for 10 assessments, dataset is analyzed based on time series models. Several time series models are selected as representatives for comparison based on Akaike information criterion (AIC),

R^2 and so forth, as shown in Table XV. For AR models, among AR (1) to AR (8), AR (6) is the best prediction model for dataset of student H in mathematics, see Table XV. AR (5) is selected to see the predicted score when the $\hat{\sigma}$ is higher than AR (6), which is almost 0. AR (1) is a more parsimonious (i.e., simpler) approach for the prediction based on past data. Based on time series models in Table XV, the accuracy of prediction of current score is depending on models based on previous past scores (p).

Table XV: Selected time series models based on dataset for student H in mathematics

Model	DF	Variance	AIC	SBC	R^2	MAE
AR (6)	3	0.0014	10.50	12.62	0.83	5.15
AR (5)	4	21.07	76.45	78.26	0.83	5.89
ARIMA (1, 1, 1)	6	395.07	84.97	85.56	0.09	15.56
ARMA (1, 1)	7	181.71	86.12	87.02	0.53	10.72
AR (1)	8	315.98	88.19	88.80	0.24	12.44

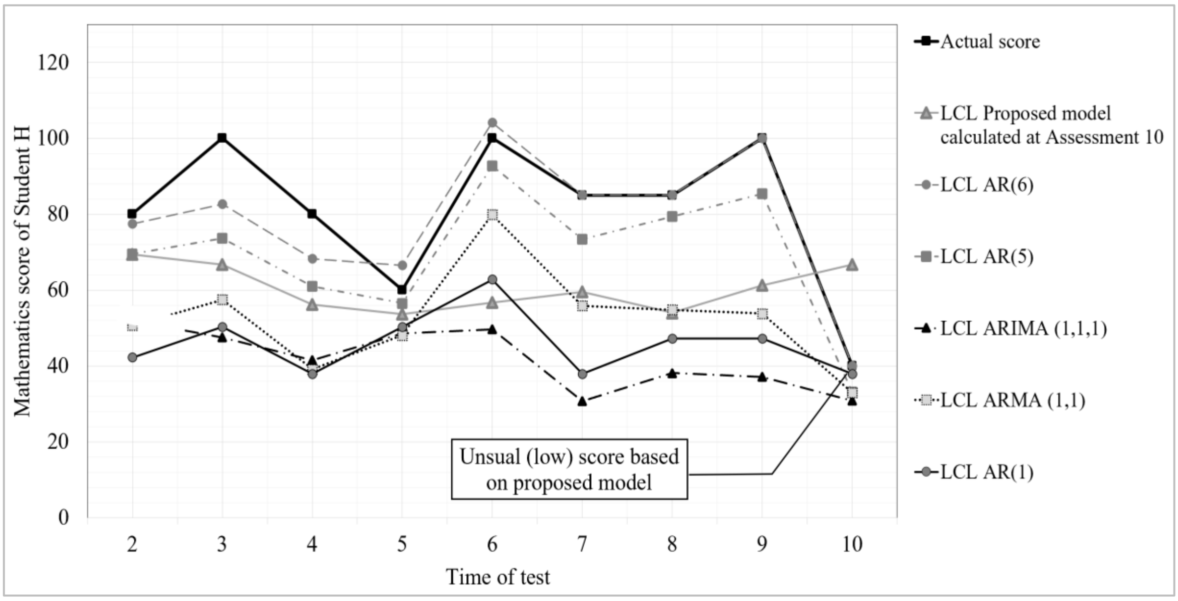


Figure XIX: Comparison of time series models and proposed model in Chapter 3 for detection of unusual scores of student H

In this analysis, the time series models are calculating an individual student score for one subject in 10 assessments, while proposed model is analyzing data of 11 students, 3 subjects, and 10 assessments simultaneously. Taking example of AR (1), the value of Y at time t is a linear function of the value of Y at time $(t-1)$ and can be expressed as:

$$Y_t = \delta + \phi Y_{(t-1)} + \varepsilon_t,$$

where Y_t is an individual student score in test t, $Y_{(t-1)}$ is an individual student score in past test $(t-1)$, δ is a constant estimate, ϕ is a lag coefficient up to order 1, and ε_t is the error term or white noise. In Figure XIX, the selected time series models do not detect the unusual score of student H, while the proposed model is able to detect unusual scores for student H. For example, AR (1) model predicted student H as $\hat{Y}_{H,1,10} = 73.43$, compared to $Y_{H,1,10} = 40$. The model considers variation of Assessment×Subject, while estimates of standard deviation includes the variation of error, which is 17.78. This $\hat{\sigma}$ is considered high to detect unusual score. The proposed model considers variation of Student, Subject, Student×Subject, and Assessment×Subject, whereas the $\hat{\sigma}$ is 10.94 and includes the error variation.

ii) For dataset analysis in Chapter 4

For the analysis, it is assumed that the current daily efforts in 2021 are equal assessments with past scores in 2020 in order to analyze by the time series models. For this analysis, it is assumed that the errors of student scores have auto correlated structure. Each individual student score in a subject for past scores of semester 1 and semester 2 examinations in 2020, six daily assessments of current semester in 2021, and a current examination in 2021 are analyzed based on time series models. Therefore, there are nine of assessments or examinations in the dataset. Table XVI shows the selected time series models for comparison. For AR model, AR (6) is selected as it is the best model to predict student scores for student M, who obtains an unexpected (low) score. Also, AR (1) is a parsimonious model for comparison.

Table XVI: Selected time series models based on dataset (example of student M) in mathematics

Model	DF	Variance	AIC	SBC	R^2	MAE
AR (6)	2	0.0009	22.17	23.55	0.34	10.00
ARIMA (1, 1, 1)	5	522.67	77.40	77.63	−0.35	19.03
AR (1)	7	412.41	81.49	882.63	0.014	15.80
ARMA (1, 1)	6	343.61	82.04	82.63	0.13	14.64

In this analysis, the time series models are utilizing an individual student score for a subject with assessments and examinations scores in sequence from 2020 to 2021. While the proposed model is using 14 students in 3 subjects, with past examinations scores in 2020 is analyzed separately with current daily assessments 2021. Figure XX shows the students score in current examination of mathematics. Through selected time series models, the unexpected (low) scores are not detected in the current examination of mathematics, while the proposed model is able to detect unexpected (low) scores for student M. Taking example of AR (1), this model predicts score of student M as $\hat{Y}_{(2021)M,2} = 58.04$, compared to $Y_{(2021)M,2} = 50$. The $\hat{\sigma}$ is 20.31, which is considered high and unable to detect unexpected (low) score. The proposed model includes variations in past scores, current daily efforts, and trend of current scores through student and subject, while the $\hat{\sigma} = 6.67$ includes only the error variation.

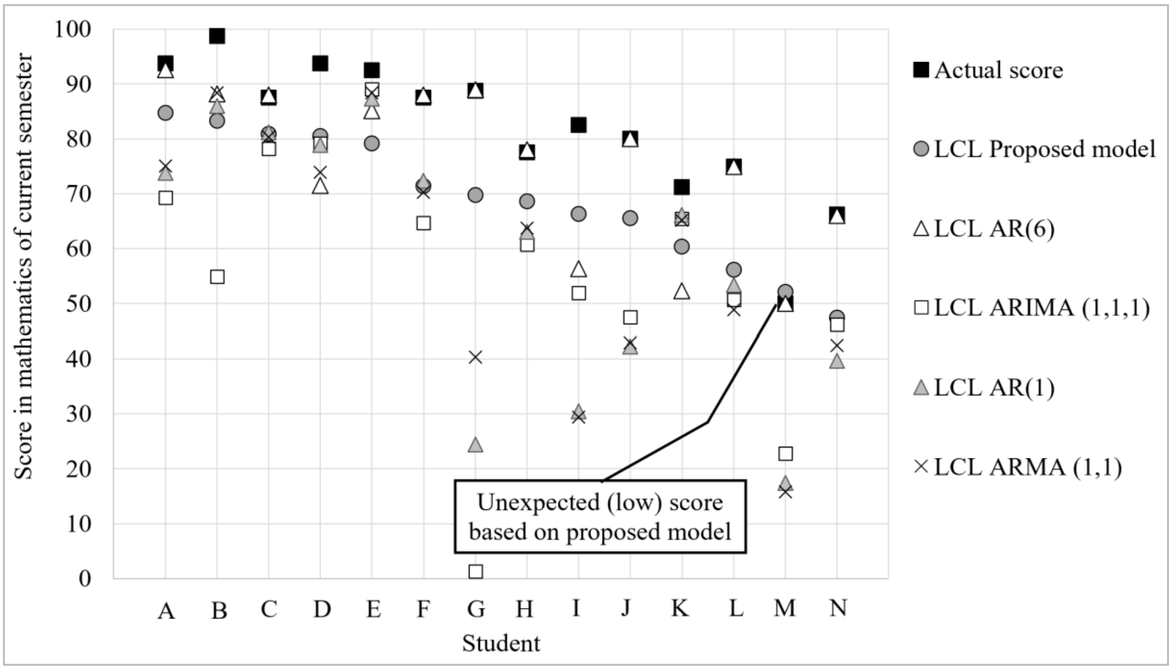


Figure XX: Comparison of time series models and proposed model in Chapter 4 for detection of unexpected (low) scores of all students in mathematics in current examination 2021

Appendix 7: Appropriate boundaries for control limits of unusual scores in education

i) Selection of k based on type I and II error

Detecting abnormality in schools depends much on the control limits as it helps to identify data that is out of a controlled state. The selection of coefficient of standard deviation for control limit is important to decide the number of abnormal data detected after an assessment. In general, the narrower the control limit, the percentage is higher that the distribution data will be recognized as out of the normal region. Data from an international school in Japan is analyzed as a case study to determine the appropriate decisions of control limit regions to assist school management in making decisions.

For the selection of coefficient k by school management, in general, it depends on the schools whether to make more detections of unusual scores through $LCL = \hat{Y}_{ijt} - \hat{\sigma}$, and less detection for $LCL = \hat{Y}_{ijt} - 3\hat{\sigma}$. Based on its definition, type I error is rejecting H_0 when it is true. The probability of committing a type I error is known as α. Whereas, type II error is failing to reject H_0 when it is false. The probability of making a type II error is labelled as β. In this research, H_0 is once the model does not detect score as an unusual score when it is a normal score. Against this statement, type I error is when in a true state, the score of an individual student is a normal score and the model mistakenly detects it as an unusual score. Students who work hard to achieve good results might get upset when they are detected as 'out of normal' although they have done their best. Also, the school will be burdened with problems that are not critical and cause exhaustion to teachers to handle the matters. On the other hand, type II error is when the score of an individual student is a truly unusual score, the model does not detect it as an unusual score. Mistaking a student who is obtaining an unusual score as normal will stop teachers from taking immediate remedies to assist the student to get back to normal.

ii) Process to determine appropriate k for control limits

Figure XXI shows the overall process for detection of the unusual scores based on daily assessments, which is described in Chapter 3. After teachers insert the students' scores from T-th daily assessments in the monitoring system, the data is analyzed based on ANOVA by using equation (3.1). Based on appropriate k selected by the school management, the unusual scores can be detected for an immediate remedy. The similar approach can be used for coefficient q in Chapter 4. For those scores which is recognized as normal scores, teachers will further monitor the results and repeat the same process of assessment, and updating students score in the school system for next assessment.

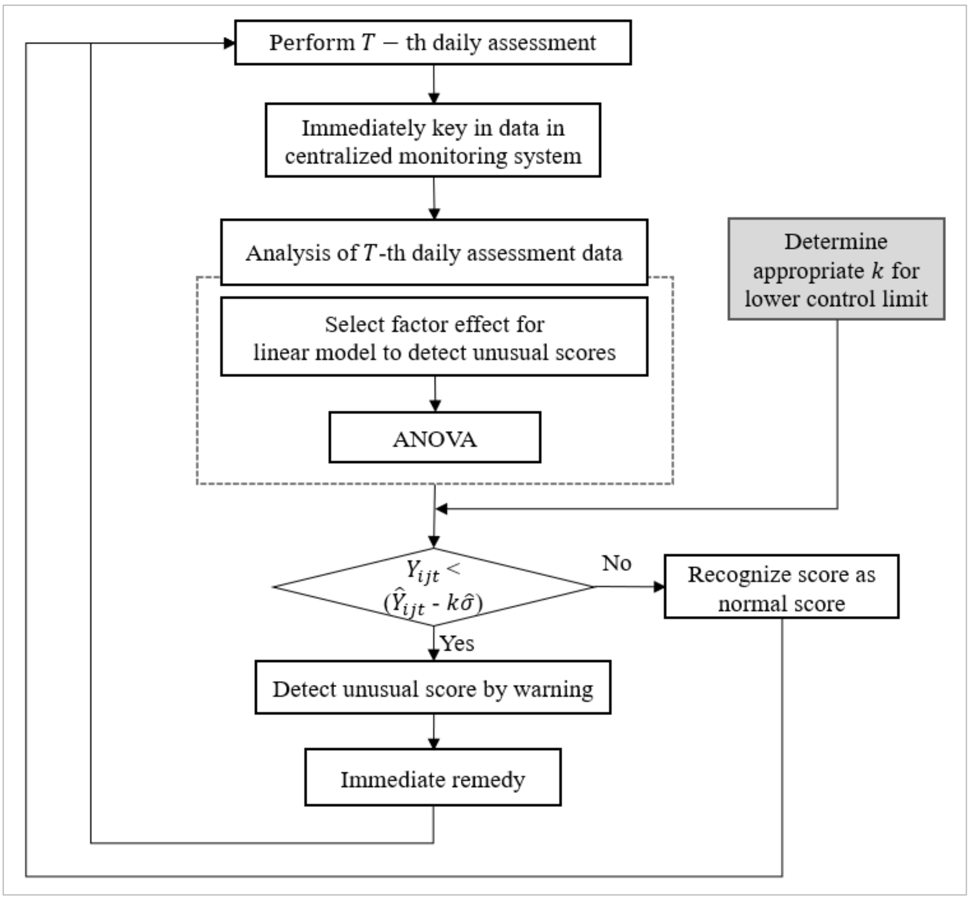


Figure XXI: Process to detect unusual scores for individual students

In general, the probability of a type I error is predetermined by the significance level. This research uses formula of P (rejecting $H_0|H_0$ true) to find α. Also, β is calculated on a μ_{TRUE}. Through formula of P (do not reject $H_0|H_0$ false), power $(1-\beta)$ is also calculated, which is the probability of rejecting H_0 when H_0 is false. This research analyzes $k = 1, 2, 3$ in relation with α and β. Finally, the case study data is used to see how it can be interpreted by the analysis of α and β.

iii) Formulation of type I and type II error

There are two types of means being used for the calculation; assumed mean, μ (under distribution of U), and $\mu_{\mathrm{TRU}E}$ (under distribution of V). The formula used is

$$\Delta = \frac{\mu_{TRUE} - \mu}{\sigma},$$

$U \sim N(0, 1^2), V \sim N(\Delta, 1^2)$. This research obtains the relationship k and α, as in Table XVII. Under σ =1, when $\Delta = 0$, $\mu_{\text{TRUE}} = \mu = 0$, $k = 1$. For this combination, type I error (α) is the region below $k = 1$. This region shows that μ is wrongly rejected when it is true. Through standard normal distribution table, z-score of -1 is 0.1587.

For the calculation of β, μ_{TRUE} is used on top of μ, when $\mu_{\text{TRUE}} = -0.5$ at $\mu = 0$ and σ=1. It can be described as $\Delta = -0.5$ ($\mu_{\text{TRUE}} = -0.5, \mu = 0$), $k = 1$, $\beta = P(V > -1)$. This region is when μ is not rejected when it is wrong. In other way, this research also calculates power, $1 - \beta = P(V < -1) = P(U < -0.5)$. These values are summarized in Table XVII. It can be seen that when k is higher, the power for detection is reducing. And when Δ is reducing, the power for detection has increased.

Table XVII: Summary of power to detect value below $\boldsymbol{k}$

Δ	$k = 1$	$k = 2$	$k = 3$
0	$\alpha = 0.1587$	$\alpha = 0.0228$	$\alpha = 0.0014$
-0.5	$1 - \beta = 0.3085$	0.0668	0.0062
-1.0	0.5	0.1587	0.0228
-1.5	0.6915	0.3085	0.0668
-2.0	0.8413	0.5	0.1587
-2.5	0.9332	0.6915	0.3085
-3.0	0.9773	0.8413	0.5

iv) Utilization of table by school management

The analysis of dataset in section 3.4 is used as an example to utilize Table XVII. From the case study, general mean is 86.74. $\hat{Y}_{H,1,10}$ is calculated as 88.61, with the estimates of error variance $\hat{\sigma}$ is 10.94. To ease calculation, let say mean score is 80 with the estimates of error variance $\hat{\sigma}$ is 10. This mean score can be mapped on Table XVII.

For example, when $\Delta = -2$, it can be said that the mean score has moved from 80 to 60. Or, in other words, when mean score has dropped from 80 to 60 for a particular time, school may see at $\Delta = -2$ row in this table. At this time, when school choose $k = 1$, it assures that 84% of the scores can be detected. This percentage is considered high and appropriate to allow school to have more detections. But if school decide to use higher k for the LCL, the percentage of detection of unusual score will reduce. If the power for detection is 50%, it means that the probability of committing type II error is also 50%.

One problem that may occur when more detections of unusual score is chosen with $k = 1$ is that the score which considered as a normal score may be detected as unusual score. When it happens, it required additional judgment for the teachers and school management to further decide whether the score is critical for immediate remedies or not. It is afraid that 'unnecessary warning' may cause teachers neglect the 'real warning'. Therefore, detecting more may not always be the best choice. For example, when a high performer student who always scores above 90 is suddenly obtained 70. Although it is good to detect that this student has dropped from his or her normal high performance, but 70 is not a critical score.

On the other hand, if schools try to avoid the above problem, probably school will choose $k = 3$ for the LCL. At this time, the detection will only be around 16%. It means that only a truly critical score will be detected as an unusual score. However, there is possibility that the unusual score, which is a truly critical score to be rescued by teacher, cannot be detected. For example, student Y is a low performer student in the cohort. Due to the score is approximately 50 in average, the estimates to produce LCL for student Y in t-th assessment is low. When student Y score below 40 and at the same time the t-th assessment is a difficult assessment, the LCL when $k = 3$ cannot detect the unusual score of student Y. Although student Y is a low performer, the score below the cut-off value (40) must be detected to enable immediate remedies. School can judge which k is suitable by judging the mean score of student data for the time and map it to this table. There is no right or wrong decision of selecting k for LCL; it depends on how school wants to manage the quality and plan for remedial measures based on resources in terms of monetary, manpower and time.

Appendix 8: Detection of unusual (high) scores based on upper control limit (UCL)

To demonstrate the detection of unusual (high) score, a dataset of 11 students, 3 subjects (mathematics, English, and science), and 8 daily assessments in a semester is selected. The results of the ANOVA table based on the model in equation (3.1) are given in Table XVIII.

Table XVIII: ANOVA table for dataset

Source	DF	Sum of Squares	Mean Square	F-ratio	Prob > F
Student: α_i	10	16118.11	1611.81	10.58	<.0001
Subject: β_j	2	719.24	359.62	2.36	0.097
Student×Subject: $(\alpha\beta)_{ij}$	20	4856.85	242.84	1.59	0.05
Assessment×Subject: $(\beta\gamma)_{jt}$	14	7697.74	549.84	3.61	<.0001
Error	217	33072.89	152		
Corrected Total	263	62464.81			

Notes: $R^2 = 0.47$, Adjusted $R^2 = 0.36$

Overall, Student, Student × Subject, and Assessment × Subject have statistically significant effects on student scores. The $R^2 = 0.47$ suggests that almost than half of the variations of student scores can be explained by their general ability, their specific ability in a subject, and a difficulty of t-th assessment for a specific subject. From the ANOVA analysis, $\hat{\sigma} = 12.34$, which is the estimate of error variance for this dataset. For the $k\hat{\sigma}$ boundary, $k = 2$ is chosen considering the balanced of type I and type II errors to indicate UCL. The $Y_{i,3,8}$, $\hat{Y}_{i,3,8}$, LCL and UCL values for each student are plotted in Figure XXI for science and it shows that Student J has obtained unusual (high) score in Assessment 8 of science. For this analysis, $\hat{\beta}_3 = -0.02$, therefore, the difficulty level of science is not remarkable. Also, student J has less interest in science, $(\widehat{\alpha\beta})_{J,3} = -6.31$ compared to other subjects and it indicates that student J might not be good or have high interest in science compared to English and mathematics. In addition, Assessment 8 of science is not a difficult assessment, $(\widehat{\beta\gamma})_{3,8} = 2.11$, compared to English and mathematics. It demonstrates students can obtain good scores in Assessment 8 of science, compared to Assessment 8 for English and mathematics.

To further investigate the reasons behind the unusual (high) scores detected in Assessment 8 for student J, this research includes the analysis of the student's previous scores from Assessment 3 to Assessment 7. The LCL and UCL for each assessment is calculated based on the scores until t-th assessment, as shown in Figure XXII. Each LCL and UCL is calculated based on t-th assessment separately. Thus, different t-th assessment analysis produce different $\hat{\sigma}$ for calculation of LCL and UCL. According to Figure XXIII, there is no detection of an unusual (high) score of student J from Assessment 3 to Assessment 7. After Assessment 8, student J has suddenly departed from the normal stable scores and has successfully obtained 95 compared to UCL, which is 92.7.

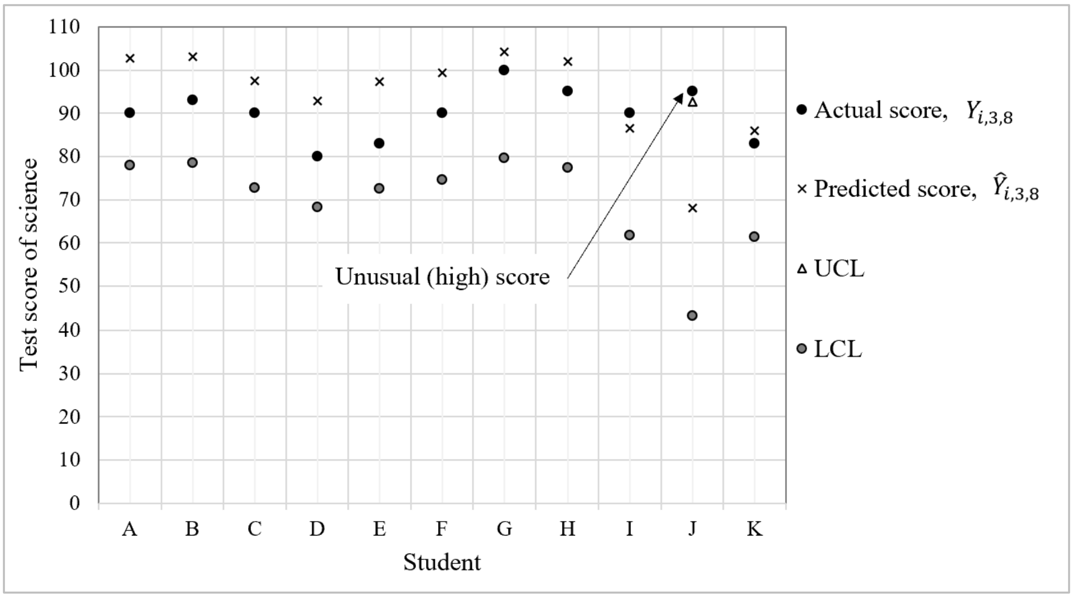


Figure XXII: Actual assessment scores against predicted scores, $\hat{\boldsymbol{Y}}_{\boldsymbol{i},\boldsymbol{3},\boldsymbol{8}}$, with LCL and UCL level of each student at Assessment 8 in science

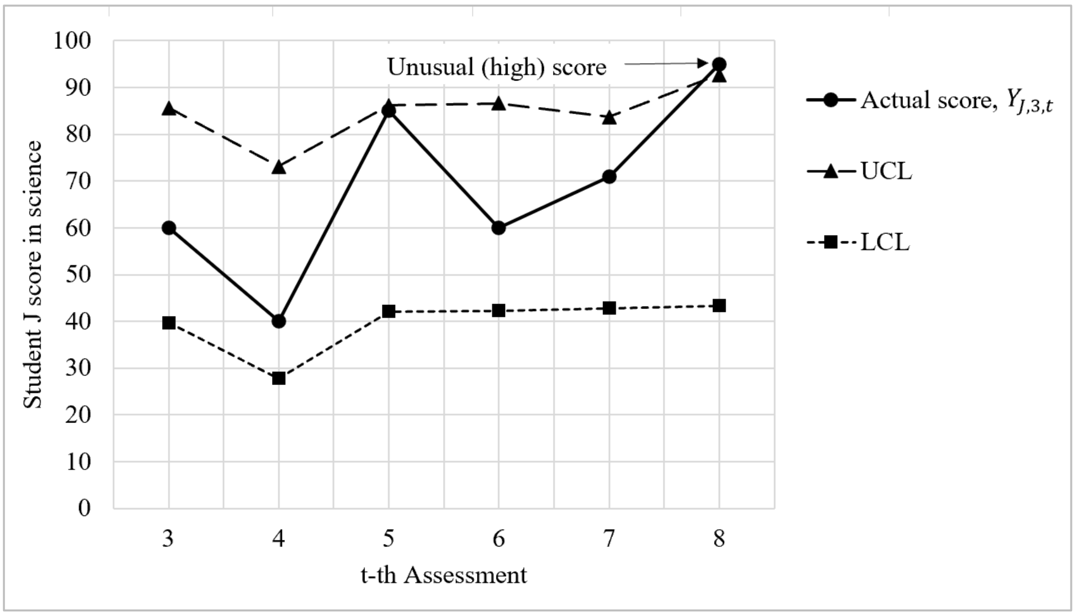


Figure XXIII: Scores of student J from Assessment 3 to Assessment 8 for science

In Assessment 8, all students scored high at more than 80. This assessment is a periodic test 2 and the students are assessed based on test for several chapters. For a teaching point of view, although all students scored well, teacher should recognize that student J has exceeded his or her expected performance. This is to boost up the motivation for the student for taking extra efforts and perform outstandingly in this assessment. Teacher also should immediately take note on what kind of teaching assistance has been given to the student other than the standard instructions provided in the daily lesson, that enables student J to get good score in this assessment. This method could be implemented to support student in future assessments as well and can be a standard to achieve good score for student J. From a learning viewpoint, student J should be proud that he or she has already achieved beyond limit and should immediately maintain the current learning style. Student J could also follow this learning styles and treat it as a standard way to achieve good score in future assessments. For the next target, student J can be more motivated to get maximum score, as it is not an impossible task if the student carries out the same momentum. For student J, this unusual (high) score can be a good practice for individual growth and as a guide in producing good score. In general, unusual (high) score could be a good practice and a benchmark for students individually and in classroom to achieve better in the next assessment.