As easy as OEE: Enabling productivity improvement in schools by using Overall Equipment Effectiveness as framework for quantitative classroom data analysis

Ilse Doyer and Wilna L. Bean

Department of Industrial and Systems Engineering,

University of Pretoria, Pretoria, South Africa

Abstract

Purpose – The purpose of this paper was to develop a quantitative classroom observation method that is able to analyse the school day to identify Time-on-Task losses comprehensively and systematically, at a level of detail that can be used by teachers and principals to stimulate and focus practical improvement efforts.

Design/methodology/approach – The novel Time-on-Task Analysis (TOTA) model was developed by triangulating the conceptual framework of the Overall Equipment Effectiveness metric with the semantics and structure of the target domain. Once developed, the model was tested structurally against a time-series classroom observation data set, after which the resulting TOTA was presented to a sample of 52 education stakeholders, who then gave their perspectives of the analysis in a structured survey.

Findings – The ontological model was found to be accurate, complete and without conceptual incongruencies, and its output novel and useful by the sample of education stakeholders. Of the participants, 90.3% found the analysis to provide a new perspective, 94.2% reported that the analysis triggered improvement ideas and 80.8%thought that their school(s) could benefit from a TOTA study.

Originality/value – The TOTA model introduces a time-loss-focused perspective to the field of quantitative classroom observation studies, which is dominated by more sociologic- and pedagogic focused topics. Its grounding in Overall Equipment Effectiveness also gives it a more detailed and systematic approach than the few Time-on-Task studies done to date, resulting in a model made for the “Gemba”: the school classroom.

Keywords Education, Productivity, School, OEE, Quantitative, Time-on-Task

1. INTRODUCTION

Hoxby (2004) rightly refers to education as “the quintessential upstream industry” as all other sectors rely on its input of educated labour. This reliance is increasing as Industry 4.0 puts higher demands on education to produce a skilled workforce able to cope with the increasing sophistication of the workplace (Spöttl and Windelband, 2021).

Apart from the important role education plays to industry, many economic models underscore a clear correlation between workforce education and a country’s economic wellbeing (Berger and Fisher, 2013, Ashton and Green, 1996). Given the centrality of the education industry, it is then concerning to note that education’s productivity has been consistently declining for decades (Creighton, 2016).

1.1 How Education Productivity is Measured

Productivity is generically calculated by the formula output divided by input. Although the formula is simple, the application and interpretation of the formula in various environments and under various circumstances is not (Schreyer and Pilat, 2001). Schreyer and Pilat (2001) argue that there are many different ways in which to measure productivity and that the metric of choice largely depends on two factors: the purpose of the productivity measurement, as well as the availability of data. They define two main categories of productivity measures:

- Single-factor productivity, such as value-added per hour worked, used in many different economic sectors such as services, manufacturing, mining, and agriculture. This type of productivity measurement thus measures an output factor in relation to a single input factor.
- Multi-factor productivity, such as gross output divided by capital-labour-energy-material-services (KLEMS), also customised for use in various economic sectors. This type of productivity measurement thus measures an output in relation to a bundle of inputs.

Much debate exists on the measurement of productivity in education for both single- and multi-factor productivity. The single factor metric most dominant in literature, and generally used at policy level, is that of learner achievement on standardised tests per dollar spent (Ahlgrim, 2010, Hoxby, 2004, Lafortune et al., 2018). As can be expected, the multi-factor productivity metrics are more stratified, depending on the input factors of interest. There is thus not one single dominant multi-factor metric available for productivity in schools (Hollingsworth and Ybarra, 2006).

When productivity growth or decline in education is discussed in literature, the dominant, single-factor, learner achievement divided by dollars spent is used for its simplicity and data availability. The problem of declining productivity in education, amidst increasing demand on schools to produce skilled labour, will not be solved through merely measuring it. The problem already confirmed by the single-factor metric – that of education's declining productivity, needs to be analysed to identify those factors that will benefit most from improvement efforts. A literature review was conducted to identify these factors and is discussed next.

1.2 How Education Productivity is Influenced

The literature on the factors influencing productivity in education is vast and includes a range of factors such as those that originate at policy level, in the home environment, at learner genetic level and within the school environment itself. Given the researchers' background of Industrial Engineering, the scope of the review was focused on those factors that fall within the 'operations' of education, i.e., at school level. From this research, two factors emerged to be consistently trumping others in terms of their impact on productivity: the quality of teaching and the amount of time the learner spends 'on task'. How these two factors were isolated, is discussed next.

In a comparative study done by Sanders et al. (1997), the effect that an individual teacher can have on a student's learning, dwarfed other factors such as class size, learner achievement level, school system and grouping of learners according to aptitude. Another study placed twins in different classes and highlighted the positive correlation between teacher experience and learner achievement (Gerritsen et al., 2017).

Perhaps surprisingly, factors such as class size, financial expenditure per student and the type of governance (private or public) have relatively little impact on learning (Ahlgrim, 2010, Dillon et al., 2002, Hoxby, 2003, Walberg, 1984). The length of a school period or school day also did not have a significant effect on learner achievement (Stallings et al., 1975, Stallings et al., 1979). Time-on-Task, the effective use of allocated academic time, was however found to have a positive correlation with learner achievement, especially in literacy and numeracy (Stallings, 1980, Fisher et al., 1978, Gettinger and Seibert, 2002, Prater, 1992).

The factors discussed here were frequently found in the literature review, but usually measured relative to a select number of other variables, making it difficult to assert that one factor is dominant as it was not compared to all factors. The literature review made it clear where school productivity improvement efforts should be focused: The place where the input of teaching is converted into the output of learning, the school classroom.

Within the school operations scope, quality of instruction and Time-on-Task were thus identified as the two potential factors that have the greatest influence on school productivity. Due to the pedagogic nature of the former factor, and the amount of literature and focus already dedicated to it, as well as the Industrial Engineering background of the researchers, the time-efficiency factor of Time-on-Task was chosen as the focus of this research.

As the scope will be on focusing on productivity within the school classroom, the story line will now backtrack slightly to consider how productivity in the classroom has been quantified to date, and then how Time-on-Task as a factor has been measured to date.

1.3 How Productivity Is Studied in The School Classroom

Wragg (2011) lists and describes the various types of classroom observation studies, which are categorised as either qualitative, or quantitative. The listed *quantitative* studies focus on personal traits, verbal and non-verbal interaction, classroom management, professional skills, teaching aids, affective, cognitive, and sociological aspects, but could include any variable of interest. The list, however, does not mention Time-on-Task as a variable commonly observed quantitatively. Even when discussing various methods of collecting data, the time-series data collection method is associated with observing when what type of interaction takes place, but not on Time-on-Task.

The qualitative methods listed focus on the same variables mentioned in the previous paragraph but allows for more flexibility in the method and variables in scope. This textbook, an introduction to classroom observations, thus focuses comprehensively on how to assess quality of teaching and other classroom aspects but omits Time-on-Task from its scope. Although Wragg (2011) notes that classroom observation studies are relatively rare, his guide on these studies indicate that where such studies do exist, they are more likely to cover different aspects of the quality of teaching than Time-on-Task.

A systematic review done by Apter et al. (2020) confirmed that the major quantitative classroom studies of the past few decades in the United Kingdom mostly focused on instructional methods, or levels and types of classroom interaction. An exception was found in the work of Stallings et al. (2014), which is discussed next.

1.4 Time-on-Task and How it is Studied in the School Classroom

The term "Time-on-Task" was first coined by Stallings (1980) , but also described earlier in the work of Carroll (1963) as "opportunity or time allowed for learning", or "time actually spent learning". In later research, Prater (1992) defined Time-on-Task as the amount of time students are engaged in academic work.

Still referred to as the seminal work on Time-on-Task (Kraft and Monti-Nussbaum, 2020), Carroll (1963) synthesised five factors that influence learning productivity into a formula (1) with the degree of learning, denoted by DL, Time actually spent by T_a and Time needed by T_n.

$$DL = f(\frac{T_a}{T_n}) \tag{1}$$

This function could be influence by three factors internal to the individual:

1. Aptitude or the amount of time needed to learn a task under optimal instructional conditions
2. Ability to understand instruction
3. Perseverance or the amount of time the learner is willing to engage actively in learning.

The factors external to the individual learner are:

4. Opportunity or time allowed for learning
5. Quality of instruction or a measure of the degree to which instruction is presented so that it will not require additional time for proficiency beyond that required in view of aptitude.

Carroll (1963) thus confirmed the two factors identified in the literature review, namely quality of teaching and Time-on-Task, as instrumental in maximising education productivity. Although this work theoretically explained Time-on-Task and its independent variables, not all the variables were possible to objectively measure in a classroom observation study. This was thus an advancement on the theory of understanding and measuring Time-on-Task, but further work would need to be done on the application of the theory to measure and analyse the productivity of a specific school.

Stallings et al. (2014) bridged this theory-application gap when specifically studying Time-on-Task by using a work sampling method to identify how classroom time was spent and to what extent learners were engaged. These Stallings Observation Studies focused on three possible Time-on-Task improvement areas: classroom management, teacher off task and lack of learner engagement. This method thus indirectly, and partially, focused on the internal factors mentioned above, and also partially, although more directly, on the fourth factor: opportunity or time allowed for learning. It is interesting to note that the method specifically excluded the more generally studied fifth factor: quality of instruction.

In correlating the variables studied by Stallings et al. (2014), with the seminal definition of Carroll (1963), a research gap was identified, which is discussed next.

1.5 The Research Gap

The Carroll (1963) model laid the groundwork in creating a theoretical understanding of Time-on-Task, but the model was not measurable, and thus not yet practical. The Stalling Observation Studies practically measured some of the Time on Task variables in the Carroll model, but not in a systematic and analytical way, omitting some measurable Time-on-Task variables from its scope, especially those relevant to the fourth Time-on-Task factor directly targeted by the studies: The amount of Time-on-

Task that is available during the school day. Unlike the other four Time-on-Task factors, this factor is both easily measured, and, more importantly, within the direct influence of school management and the teacher, and thus a good starting point for influencing classroom productivity. The analysis of the existing work on Time-on-Task thus identified a research gap, formulated into the following research question:

How can Time-on-Task in the school day be systematically, and comprehensively, analysed to enable teachers and school management to improve their school's productivity?

No further practical, analytical Time-on-Task models were found in the literature, but the themes within, and terminology of Time-on-Task in literature lead the researchers to identify a potentially novel way of processing school day data. With their Industrial Engineering background, the researchers identified that much of the Time-on-Task terminology used in literature, such as 'allocated academic time', and 'teacher loading time' corresponded conceptually to terminology found in a diagnostic productivity metric used in Total Productive Maintenance, a branch of Lean Management used mostly in manufacturing, namely Overall Equipment Effectiveness or OEE (Gibbons and Burgess, 2010).

1.6 The Research Contribution

Using OEE as theoretical framework, this research developed the novel Time-on-Task Analysis (TOTA) model with which to analyse quantitatively and systematically, in more detail than what has been applied to date, how the school day is spent and what specific Time-on-Task losses occur. The systematic and detailed analysis of Time-on-Task losses over the course of a complete school day, provides a unique and diagnostic perspective, that can be used at a practical level by teachers and principals to target productivity improvement efforts.

Additionally, with quantitative classroom observation studies being relatively rare (Wragg, 2011), and where done, mostly focus on pedagogic or sociological factors (Apter et al., 2020, Wragg, 2011), the TOTA model enables time-loss focused quantitative studies, at a practical, grass roots level.

Although the research was initiated by a request from a large private schools' group, and thus relevant to the current education sector, scientific rigour was assured by grounding the model in existing theory. This theoretical framework is discussed in section 2, after which the research methodology will be discussed. A results section is then followed by a conclusion with further research opportunities listed.

2. THEORETICAL FRAMEWORK

In 1988, Nakajima introduced a novel quantitative metric for the measurement of productivity, namely OEE, to increase the productivity of machines and equipment (De Ron and Rooda, 2006). OEE expresses the productivity of a machine as a function of its availability, performance, and quality rates. As can be seen in Figure 1, these three elements are further broken down into what is known as the 'the six big losses' (Muchiri and Pintelon, 2008).

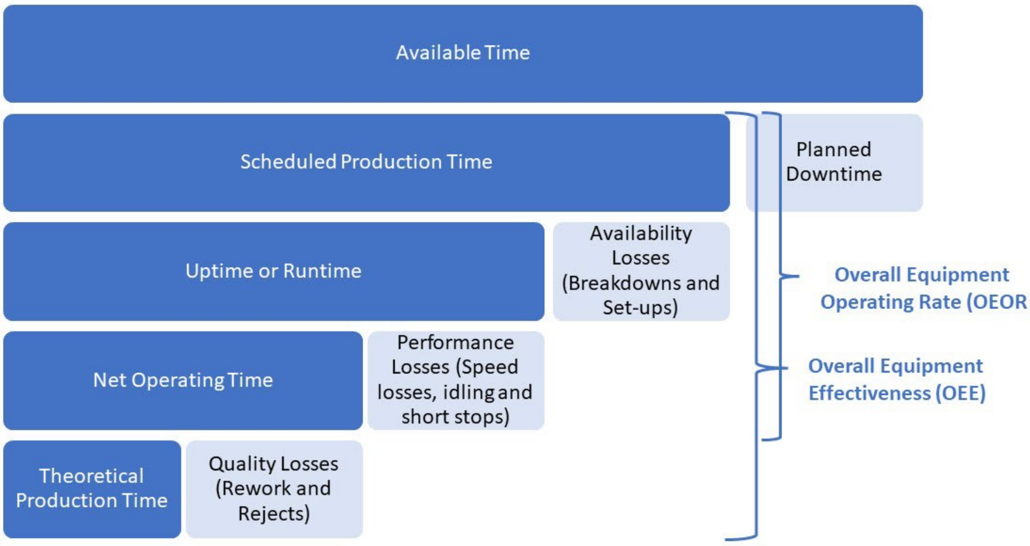

Figure 1: Overall Equipment Effectiveness Structure

Although mostly applied in the manufacturing sector, OEE has been adapted to analyse operating losses in urban freight transportation systems (Muñoz-Villamizar et al., 2018) and (repair) service organisations (Facchinetti and Citterio, 2022). Butlewski et al. (2017) adopted OEE to create Overall Labour Effectiveness with a specific focus on worker safety and fatigue. Such adaptations are scarce in literature and differ greatly in what and how they chose to translate OEE but were useful to inform how such a model could be adapted in different ways.

Literature was also screened, and experts from different fields consulted, for alternative metrics to OEE. Only one other widely used metric was found that has the same systematic and analytical approach as OEE: Du Pont Analysis, used in Financial Accounting (Jansen et al., 2012). Although structurally similar to OEE, and thus equally powerful in its diagnostic ability, Du Pont Analysis was contextually equally far removed from the school classroom, and thus did not present a clear advantage. Additionally, literature on the theory as well as application of OEE was also more ubiquitous, and OEE was thus selected as theoretical framework.

The strength of the OEE metric lies partly in its multiplier effect, which prevents various productivity factors from being measured in isolation, or hidden by aggregation, but rather highlights the compound effect that these factors have on productivity overall. This is demonstrated in the formula for OEE, as described in the seminal work of Hansen (2002).

$$OEE = A^R \, x \, P^R \, x \, Q^R \tag{2}$$

Where A^R denotes the Availability Rate, P^R the Performance Rate and Q^R the Quality rate. Where Operating Time is denoted by OT, Loading Time by LT, Net Operating Time by OT_n and Valuable Operating Time by OT_v, the expanded form of OEE is expressed as:

$$OEE = \frac{OT}{LT} \, x \, \frac{OT_n}{OT} \, x \, \frac{OT_v}{OT_n} \tag{3}$$

Therefore, essentially reducing to the compound effect:

$$OEE = \frac{OT_v}{LT} \quad (4)$$

Which correlates with the generic formula for productivity (*P*) of Output (*O*) divided by Input (*I*):

$$P = \frac{O}{I} \quad (5)$$

Despite its ability to be reduced to a single-factor productivity metric, OEE is not predominantly a productivity measurement metric. Its strength lies in its multi-factor capabilities to diagnose – pinpoint and quantify - productivity losses.

The concept of Time-on-Task however is the equivalent of Net Operating Time in the OEE formula and thus more closely resembles what is measured by a truncated, lesser-known version of OEE, called Overall Equipment Operating Rate (OEOR), which does not take into account the quality aspect included in OEE (Kobayashi, 1995). Figure 1 demonstrates the difference between OEE and OEOR. Little published research exists on OEOR, hence OEE remained as the theoretical framework for the research.

When the TOTA model was validated using a set of classroom observation data (discussed in section 4.3), the data confirmed the choice of using the truncated version of OEE as this data set contained data that could be easily observed during a time-series study done in a classroom setting. The reason for this is that *effective* learning cannot be observed during classroom observation studies but will require some form of assessment of learning absorption and retention. Classroom observation studies can however be set up to easily assess the availability and much of the performance data of a group of learners. How the truncated version of OEE was used as a conceptual framework is described next.

Five derivatives of OEE were identified in literature and critically assessed as theoretical frameworks as alternatives to OEE. Total Effective Equipment Performance (TEEP) as proposed by Ivancic (1998) is structured to also explore the causation between planned downtime and unplanned downtime with a specific focus on the effects of planned maintenance. In Production Equipment Effectiveness (PEE), Raouf (1994) postulates that the three elements of OEE should be given a weighting instead of automatically being considered of equal weight. Both of these derivatives of OEE make valuable contributions but are still derivatives of the generic framework. For the purposes of creating a basic model for analysis, the generic OEE was considered a better starting point for development, and these two derivatives possible refinements at a stage where the basic model has been established and used to the extent where the weightings and causative relationships become clearer. The other three derivatives, described in a review by Muchiri and Pintelon (2008), all focus at an organisational, instead of machine level, meaning the data is more aggregated. As this research aims to propose a practical model that enables improvement through detailed analysis, none of these derivatives posed a valid challenge to OEE as theoretical framework.

3. RESEARCH METHODOLOGY

The aim of the research was to develop a quantitative classroom observation method that is able to analyse the school day to identify Time-on-Task losses comprehensively and systematically, at a level of detail that can be used by teachers and principals to stimulate and focus practical improvement efforts. Building on the theoretical work of Carroll (1963) and the more practical work of Stallings et al. (2014), the research question is: *How can Time-on-Task in the school day be systematically, and* *comprehensively, analysed to enable teachers and school management to improve their school's productivity?*

With its focus on improvement of the research target domain by introducing novel and innovative design artifacts (Simon, 1996), Design Science Research (DSR) was chosen as the research paradigm, using the three-cycle framework of Hevner (2007), consisting of:

- The Relevance Cycle, which connects the target environment with the design activities
- The Rigor Cycle, which ensures that the design activities are grounded in established theory
- The Design Cycle, which iteratively builds and evaluates the design artifact by interacting with both the Relevance and Rigor Cycles.

This research was triggered in the Relevance Cycle by the request of the Chief Executive Officer of a large private schools' group to find out how productivity could be measured and improved at school level. The research into the topic of productivity in schools was set out in section 1 of this paper – this formed part of the Relevance Cycle, where the researchers familiarised themselves with the vocabulary, dynamics, and metrics of the target domain. It was during this phase that the Rigor Cycle was set in motion when the potential of using OEE as theoretical framework was identified (discussed in section 2). OEE, as a well-established, state-of-the-art diagnostic productivity metric could provide a grounding theoretical base from which to develop a similar productivity diagnostic tool for schools, during the iterative Design Cycle.

The Design Cycle, of artifact development and testing, was thus done by triangulating the existing and proven theoretical base, namely OEE, with terminology and definitions from the target domain of education. Once designed, the conceptually translated diagnostic model was tested by using the model to code a set of time-series school day observation data to check for accuracy, completeness, and freedom of conflict (Shanks et al., 2003) against the semantics and structure of the target domain. Once these iterations of structural testing had created a structurally robust model (another round of the Rigor Cycle), the model was tested for its usability, novelty, and value by processing quantitative classroom observation data to produce a Time-on-task school day analysis (a round of the Relevance Cycle). This model output was then presented to focus groups of education stakeholders, after which their perspectives on the novelty and utility of the TOTA model output was collected through a structured survey (another round of the Relevance Cycle).

Table 1 lists DSR requirements developed by Hevner et al. (2004) for the Information Systems environment, but as paraphrased by Venable (2010) for more generic use in various DSR environments. The final column in the table indicates to what extent the requirements were met by this research.

Table 1: DSR Requirements Evaluation of this Research

DSR requirement (Hevner et al., 2004)	Description (Venable, 2010)	Evaluation of this research
1. Design as an Artifact	An identifiable and viable design artifact must be produced.	The TOTA model was produced (section 4.1)
2. Problem Relevance	The design must address a relevant and important problem.	The research was triggered at the request of a large private schools' group to study and improve productivity at school level.

		A sample group of 52 education stakeholders mostly from the commissioning schools' group, overwhelmingly found the result of the TOTA model to be useful and novel (Chapter 5).
3. Design Evaluation	The utility, quality, and efficacy of the design artifact must be rigorously evaluated.	The quality of the model was evaluated against qualitative data from the target domain (section ???). The utility was assessed through focus groups of stakeholders from the target domain. The efficacy of the model to *energise* productivity improvement was established (section ???), but the translation of this created intention into measurable productivity improvements is still to be tested through longitudinal studies (section 6.2).
4. Research Contributions	The contribution must be clear and verifiable. Contributions include the design artifacts themselves, new foundations, and new methodologies.	The TOTA model has been developed into more detail than its grounding theoretical framework (Chapter 4), along with guidelines for implementation (sections 4.2 and 6.1).
5. Research Rigour	Research methods must be rigorously applied.	The use of the three-cycle DSR approach and the evaluation shown in this table, ensured rigour in the research.
6. Design as a Search Process	Research must be conducted with knowledge of, and through iterative testing with other, competing approaches.	Two Time-on-Task models were evaluated and used as basis to identify the research gap (section 2.4). Other quantitative studies were studied but did not deal with Time-on-Task (section 2.4). The OEE framework was used iteratively throughout the development of the TOTA model (Chapter 2).
7. Communication of the Research	Presentation of results needs to address both the rigour requirements of the academic audience and the relevance requirements of the professional (e.g., managerial) audience.	The research was conducted in successful fulfilment of a Masters' Degree in Industrial Engineering at the University of Pretoria, South Africa, of which this peer-reviewed article is the publication. Additionally, a sample of 52 education stakeholders included academics, corporate managers, principals, and teachers found the research output to be novel and useful (Chapter 6)

The DSR requirements 1,2,4,6 and 7 were thus fulfilled to a high extent. Requirement 3 had a good complementing model (OEE) to inform the research, but no models were conceptually close enough to truly compete. Requirement 5 was partially fulfilled and should thus be the task of further research. The finer detail of how the research was conducted will be presented alongside its results, which is discussed next.

4. RESULT OF THE DESIGN CYCLE: THE NOVEL TOTA MODEL

The development of the TOTA model first focused on conceptually translating the high-level structure of the truncated OEE (Figure 1) into the vocabulary of the target domain (Figure 2), using the terminology and definitions identified during the literature review of the Relevance Cycle.

4.1 STRUCTURAL PRESENTATION OF THE TOTA MODEL

Figure 2 visualises the basic concept of the TOTA model: first availability, then Time-on-Task losses systematically reduce the scheduled academic time. When compared to Figure 1, Figure 2 also shows the structural overlap with OEE. Being structurally similar to OEE, the TOTA framework therefore provides a means of systematically identifying and prioritising Time-on-Task losses in the school day.

Where Stallings et al. (2014) focused on three variables, namely classroom management, teacher off task and lack of learner engagement, the TOTA model focuses on a systematic analysis of all time-losses within the school day.

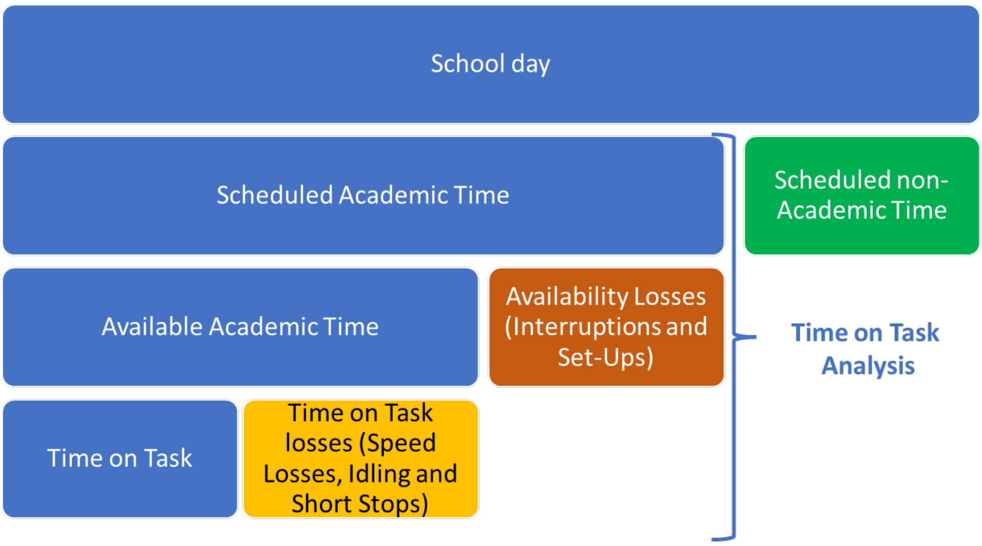

Figure 2: Time-on-Task Analysis Structure

Breaking these loss categories into more detail (Figure 3) enables the quantification and identification of detail loss categories to be prioritised in production improvement efforts to a level of detail recognisable to, and within reach of, teachers and principals.

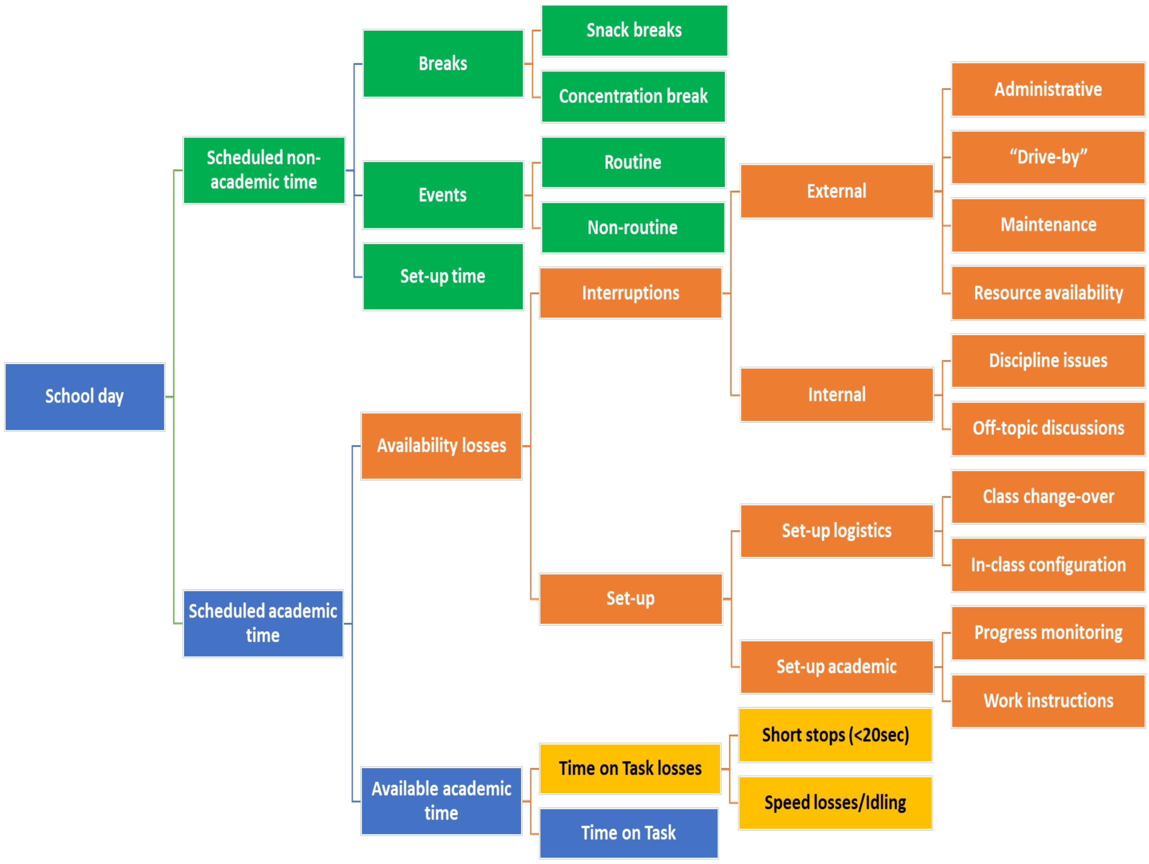

Figure 3: The TOTA Model Showing Detailed Time-on-Task Losses in the School Day

As indicated in Figure 3, the TOTA model first separates scheduled academic and non-academic time. In OEE, scheduled downtime does not form part of the OEE calculation, as scheduled downtime is a utilisation issue, the control over which lies with management rather than the operations team responsible for the productivity of the machine being measured. Similarly, in the classroom, scheduled non-academic events such as school assemblies or school rosters requiring class changeovers, fall outside the control of the classroom teacher, responsible for optimising Time-on-Task otherwise.

Set-up time is allocated twice in the detailed model of Figure 3, appearing to be a duplication. As will be explained in the detail definitions of each loss category in the section 3.2, the *scheduled* set-up time is the optimal time reserved for set-up, for example a target class changeover time set as three minutes. Any time taken *more than the scheduled* time, is considered an availability, not a utilisation, loss. This practice mirrors the treatment of change-overs in OEE (Hansen, 2002).

The scheduled academic time is separated into availability losses, such as interruptions and non-scheduled set-up-time, and available academic time. There is thus scheduled time, when learners are supposed to be taught in the classrooms, but also some availability losses such as interruptions and extra set-up time which consume some this schedule academic time, leaving the category 'available academic time'.

Some of this available academic time might be consumed by further productivity losses, such as Time-on-Task losses, which take place when learners are supposed to be on task but have lost concentration (short stops in OEE) or are waiting for others to finish (idling or speed losses in OEE).

What remains after these time losses have been removed from the school day, is Time-on-Task: The moments spent on academic activity. This Time-on-Task (T_{oT}) could thus be a form of output measure for a productivity measure of the scheduled academic time (T_{SA}) as the input measure of the teacher and learner-controlled part of a school day. The productivity could be expressed as a Time-on-Task ratio (T_{oT}) as follows:

$$T_{oT}\ \% = \frac{T_{oT}}{T_{SA}} \tag{6}$$

The point of metrics such as OEE and the TOTA model is not to determine that there are Time-on-Task losses, but rather where these losses occur, and which ones are most acute. That is their real contribution: the focusing of the improvement effort. The expanded, analytical form of the Time-on-Task ratio is thus of more interest to this research:

$$T_{oT}\ \% = \frac{T_{AA}}{T_{SA}}\ x\ \frac{T_{oT}}{T_{AA}} \tag{7}$$

Where T_{AA} denotes Available Academic Time.

The Availability ratio (A^R) of OEE is thus represented in the TOTA model as:

$$A^R = \frac{T_{AA}}{T_{SA}} \tag{8}$$

This formula thus expresses the Availability ratio as an efficiency ratio of available time over scheduled time, thus accounting for availability losses. The Performance ratio (P^R) of OEE is represented (to a lesser extent) in the TOTA model as:

$$P^R = \frac{T_{oT}}{T_{AA}} \tag{9}$$

This ratio thus accounts for performance losses by calculating the ratio of Time on Task to Available time. The phrase 'to a lesser extent' is used here since speed loss cannot as yet be individually measured per student. In the classroom observations it was visible when and how many learners were idle, but their speed of working could not be compared to their ideal speed, as OEE requires it to be measured (Hansen, 2002). Nevertheless, this Performance measure still brings the productivity analysis closer to Time-on-Task analysis.

As shown in Figure 3, the TOTA model consists of six hierarchical levels of detail, with the conceptual equivalent of the six big losses of OEE found at the fourth level. Two more layers of detail than contained in the structure of OEE (Figure 1) were thus added to further describe and organise Time-on-Task losses to enable improvement.

4.2 PRACTICAL DEFINITIONS OF THE TOTA LOSS CATEGORIES

To ensure that the TOTA model does not remain only a theoretical concept, the detail definitions of each of the categories were developed to make the theory practical and the model accessible to school staff and productivity improvement practitioners alike. These definitions were developed by studying the detail OEE definitions in the seminal work of Hansen (2002) and finding the conceptually equivalent terminology in education literature. As the contexts of the two models are different, not all of the OEE terminology could be found in education literature, and vice versa, and the researchers thus often had to conceptually structure and define the TOTA model and its categories based on the knowledge gained in the literature reviews, and from their knowledge of the target domain. This corresponds with the recommendation of Hevner (2007) not to let the Rigor Cycle become a constraining, rather than a stimulating, input. The detailed definitions developed during the DSR, are presented here in Table 2.

Table 2: TOTA Model Category Definitions

School Day	The official school hours when learners are all required to attend and includes both scheduled academic time and scheduled non-academic time.
Scheduled Non-Academic Time	All minutes of the school day scheduled to not be used for academic activities. This may include: • Scheduled breaks including scheduled snack breaks • Scheduled concentration breaks, e.g., between periods or tasks to enable sustained concentration during the scheduled academic time. • Routine or non-routine scheduled non-academic events approved by school management, e.g., school assemblies, sports team send-offs, camps etc. • Optimal set-up time. The optimal set-up time deemed necessary to do set-up activities, described under availability losses. Since set-up activities are mostly necessary to deliver optimal academic performance, set-up activities should not be deemed as availability losses in their entirety, but should be encouraged by allocating an optimal, or ideal. amount of time as scheduled. In keeping with OEE (Hansen, 2002), all scheduled non-academic activities should be categorised into optimal time and loss time,e.g., if the optimal meeting time is 30 minutes, but the meeting takes 45 minutes, the first 30 minutes are deemed scheduled set-up time and the last 15 minutes are considered a set-up loss. The ‘optimal time’ should be reviewed as efficiency within the system improves.
Scheduled Academic Time	All minutes of the school day scheduled to be used for academic activities, whether it was used eventually or not. In education literature, this term is described as “opportunity” (Carroll, 1963), or as “allocated academic learning time” (Fisher et al., 1978), or as he upper limit of in-class opportunities for students to be engaged in learning (Gettinger and Seibert, 2002). Although similar, the scheduled academic time proposed in this study accounts for scheduled losses differently and hence the new term as proposed here.
Availability Losses	All minutes that were scheduled for academic activities but used for non-academic activities instead. These could include straight-forward “downtime” losses such as time lost due to a projector not working when it was needed, or the excessive set-up times explained under Scheduled Non-Academic Time above. A third subcategory is interruptions, split into internal and external interruptions in education literature (Fisher et al., 1978, Kraft and Monti-Nussbaum, 2020). Note that, borrowing theoretically from the excellent definition of change-over time by (Shingo, 1996): Any stops in academic time should be timed from the moment the learner stops being fully on task to the moment the learner is fully on task again. Thus, the interruption does not cease when the interrupter leaves the room or the teacher asks the learners to return to their activities, it ceases when the learners are fully on task again.
Interruptions	Interruptions are unplanned incidents during scheduled academic time that prevent academic work from being done and last for more than 20 seconds. These interruptions can be internally or externally induced.
External Interruptions	This category of unplanned interruption of more than 20 seconds comes from outside of the classroom that are not under the direct control of the class teacher (Kraft and Monti-Nussbaum, 2020), e.g., an intercom announcement, a delivery,

	or someone requesting audience with the teacher or class. Some further sub0-categories have been defined: • Administrative - Interruptions initiated by school administration or management such as announcements, messages, class callouts of teacher or learners, etc. • "Drive-by" (Kraft and Monti-Nussbaum, 2020) - Interruptions initiated by fellow staff members or learners such as meetings between teachers in class time, phone calls taken from parents or others, other classes making noise that prevents further instruction, etc. • Maintenance - Interruptions caused by the failure or unplanned maintenance of infrastructure or resources such as loud noises e.g., lawn mower, burst pipes / broken window, faulty equipment e.g., projector, tablets etc. • Resource Availability - Interruptions caused by the unavailability of facilities or learning resources such as double bookings, unavailability of rooms or equipment, missing resources, etc.
Internal Interruptions	Interruption of more than 20 seconds that come from inside the classroom. Although literature refers to internal interruptions as being caused by off-task student behaviour (Little and Akin-Little, 2008), the definitions here also recognise that teachers can also be the cause of interruptions. Two sub-categories have been defined: • Discipline Issues - Interruptions caused by, or being the result of, poor behaviour such as disruptive behaviour by learner(s), learners ignoring instructions, teachers excessively reprimanding in terms of time or target audience, etc. • Off-topic Discussions - Interruptions to the academic task at hand caused by off-task discussions – academic or non-academic – that cause a disruption in concentration or a delay in the completion of a task. These include teacher-initiated interruptions such as announcements or off-task questions while students are completing a workbook activity etc. or learner initiated interruptions such as casual conversation, asking that instructions be repeated where instructions were clear or asking to borrow stationery etc.
Set-up	Activities that enable learning or teaching activities but are not in themselves academic learning for learners. IMPORTANT: Only the minutes spent on set-up activities that exceed the optimal scheduled time for these activities are included here. See the discussion on Scheduled non-academic time for how this is to be determined.
Logistics	Activities that mostly involve movement, e.g., moving learners or teachers between places of learning, moving learning resources into position, or transportation. Its subcategories are class changeovers and in-class configuration.
Class Changeovers	Movement between places of learning, e.g. learners moving from the laboratory to the English class. This activity starts the moment learning stops to the moment learning commences again.
In-class configurations	Non-learning or -teaching related actions during class time with the intention of enabling further academic activities, e.g., projector set-up, handing out teaching material, separating, or grouping desks for activities, taking out or putting away books or stationery, etc.

Set-up: Academic	Actions related directly to setting up for academic work but are not the teaching or learning of content in themselves. Its subcategories are progress monitoring, work instructions, progress monitoring, and assessment marking.
Progress Monitoring	Actions taken by the teacher to review progress of academic work such as teacher signing workbooks, teacher asking how far learners are on an assignment or how many have finished an activity, teacher checking whether learners have done what was required, teacher asking the class what work has been covered, etc.
Work Instructions	Actions taken to give instructions on academic work such as giving or receiving instructions regarding academic work to be done, coordination of groups for assignments, learners asking follow-up questions on instructions, etc.
Assessment Marking	Actions taken to mark assessments, capture those marks on a system and give feedback to learners regarding the mark only. Detailed feedback or discussion on the assessment is classified as revision.
Available Academic Time	All scheduled academic time left for academic work in between set-ups and interruptions. Although interruptions and inefficient setups have now been accounted for, there might still be short stops and idle time losses built into what we consider to be time left over for academic learning.
Time-on-Task Losses	Time-on-Task losses refer to those seconds lost due to short stops of less than 20 seconds or idle time, defined below.
Short Stops	Any unplanned interruptions and excessive set-ups, as defined above, of less than 20 seconds.
Idling / Speed Loss	The difference between the theoretical time that an activity should take and the actual time it took to do an activity (Hansen, 2002). A learner taking longer to do an activity than what it should take, is an efficiency loss in the classical sense. When learners finish a test or activity sooner than the rest of the class and are expected to wait idly, or keep themselves busy with non-academic activities, constitutes an idling loss. When work is explained, but not understood, or repeated when already understood, also constitute idling losses as the brain is running, or disconnected, but not doing useful work (Merriam-Webster, 2002).
Time-on-Task	Time spent paying attention, or trying to learn (Stallings, 1980, Carroll, 1963). The amount of time students are engaged in academic work (Prater, 1992).

During the Design Cycle, the TOTA Model was thus developed iteratively by interacting with the structure and semantics found in the literature of the target domain (Relevance Cycle), but also that of the theoretical framework used (Rigour Cycle).

4.3 MODEL VALIDATION AT AN ONTOLOGICAL LEVEL

Before doing field testing, the theoretical model was tested structurally by using the ontological approach by Shanks et al. (2003) who reason that the validation of a conceptual model involves testing its faithful representation of the domain it is intended to represent. It is regarded as faithful when the model is accurate, complete, and conflict-free through the perspective of the domain stakeholders. Shanks et al. (2003) provide the following definitions for each attribute:

- Accuracy: The model should accurately represent the semantics of the focal domain as perceived by the stakeholder(s) of the focal domain.
- Completeness: The model should completely represent the semantics of the focal domain as perceived by the focal stakeholder(s)
- Conflict-free: The semantics represented in different parts of the model should not contradict one another; and there should be no redundancy. To reduce the likelihood of conflicts arising

if and when the model is subsequently updated the model should not contain redundant semantics.

These validity attributes were tested by coding the semantics and structure of the focal domain into a template created from the TOTA model. The semantics and structure were sourced from a secondary data set of classroom observations made available to the researchers by the private schools group mentioned earlier. The data had been collected during a time-series, systematic, quantitative classroom observation study in the intermediate-senior phase of one of its primary schools in February 2020. The data set contained 450 activities undertaken during 44 school periods and contained both qualitative data (descriptions of activities, as well as field notes) and quantitative data (start times and durations of activities).

This first stage of validation found that up to the fourth hierarchical level of the TOTA model (Figure 3), the model proved to be accurate, complete, and conflict-free with no need for any additions, omissions, or adjustments.

At the two detail levels, where four of the six major losses had been broken into two additional layers of detail, only seven (or 1.5%) of the 450 activities were categorised as “other” as they did not fit into any of the 32 TOTA categories. On closer examination, these activities were ad hoc, such as a swimming team send-off ceremony, to justify the creation of an additional TOTA category and it was thus concluded that the model could accommodate such ad hoc occurrences in the “other” categories provided under each of the major loss categories. These observations thus did not challenge the validity of the structure or semantics of the model.

The detail coding of the classroom data did, however, provide opportunity to strengthen the detail definitions of the TOTA categories by providing additional practical examples of the loss categories, making the description of the TOTA model more robust.

The concept of conflict freedom screens for redundancy and contradictions in a model (Shanks et al., 2003). Although no contradictions were found, two of the sixth (most detailed) level categories could be collapsed, and one other such category expanded to clarify their use or to make the model more streamlined.

At an ontological level, the TOTA model was thus found to be valid by testing it against the realities, structure, and semantics of the focal domain. This meant that the model was ready to be tested in the field, which is described next.

5. RESULT OF THE RELEVANCE CYCLE: AN ENERGISED TEACHER CORPS

As stated by Hevner (2007), the DSR output must be returned into the target domain for evaluation. The TOTA model was used to descriptively analyse the durations of school day activities, i.e. the quantitative part of the 450-line classroom observation data set described earlier. This data provided a rich source for analysis, which produced a wide range of graphs, presented in a focus group setting to various education stakeholders.

Seven focus groups were held with a total of 52 education stakeholders participating. The breakdown of the stakeholder sample group was as follows: 39 teachers, 2 senior academics, 4 specialists, 4 principals or Heads of Department and 3 corporate managers. The participating teachers and one principal were recruited from the school where the classroom observation study had been done and the other participants were recruited from the large corporate education group which the participating school formed part of.

The work of Brown and Jayakody (2008), which builds on that of Shanks et al. (2003), as well as that of DeLone and McLean (1992), was used to identify another four validation attributes to be tested in structured surveys given to the participants after the Time-on-Task Analysis presentation. Table 3 shows these attributes and the related survey statements with Figure 4 showing the results of the survey.

Table 3: Validity Attributes Tested with Likert-scale Survey Statements

Validity Attributes (Brown and Jayakody, 2008)	Survey statement (Likert scale)
Individual impact	"The analysis provides a new perspective on how productivity can be studied in the classroom"
Individual impact Perceived usefulness	"I came up with improvement ideas during/after the presentation"
Perceived organisational impact Perceived usefulness Intention to use	"The school/s I am involved in can/could benefit from such a classroom work study project"

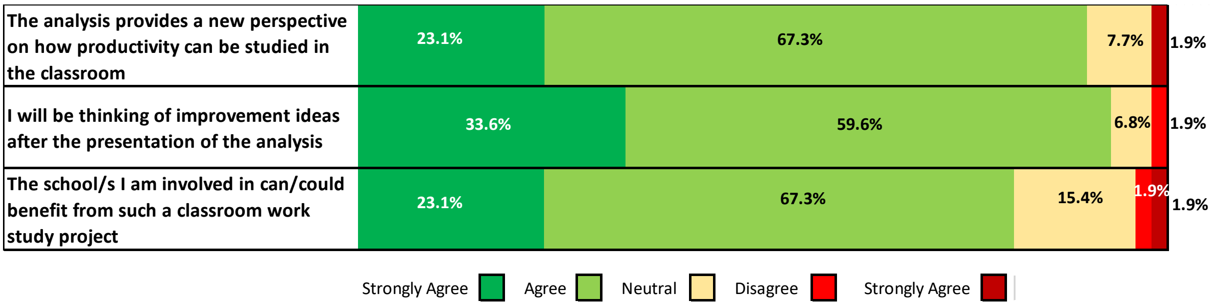

Figure 4: Education Stakeholders' Responses to Survey Statements

The Likert-scale data were analysed using both descriptive and inferential statistics (Heeringa et al., 2017, Boone and Boone, 2012), with t-tests used to identify differences in responses between the group consisting of teachers, and the group consisting of mostly managers.

From Figure 4, it is clear that the output of the TOTA model was well received by the target domain, with 90.3% of the 52 education stakeholders sample responding with either agree or strongly agree to the novelty of the TOTA model outputs, and 80.8% to the utility attribute, satisfying the Hevner et al. (2004) requirement that the output of the model be presented to the users as well as managers of the target domain for testing.

Given the fact that literature generally describes educators as being notoriously resistant to change (Elmore, 2004, Senge, 1995, Sarason, 1990), the second and third statements showed that the TOTA model had achieved the opposite: 94.2% of the focus group said the TOTA model output had mobilised them toward improvement thinking, and 90.4% would welcome a TOTA study in their school(s). T-tests of these statement's data revealed that, with a mean of 4.03 (with 4 = agree, 5 = strongly agree), the teachers were slightly more enthusiastic about such a study in their school than were the managers with a mean of 3.85. The "improvement ideas" statement showed slightly more momentum on the part of the managers and academics with a mean of 4.38, as opposed to that of the teachers at 4.29, but indicating that both teachers and managers had been energised toward productivity improvement.

6. CONCLUSION

Despite the central and upstream role schools play to the economy, productivity in schools has not kept up to the productivity improvements in other sectors (Creighton, 2016). Vast amounts of literature focus on the factors of productivity in education, but Time-on-Task, a dominant factor in school productivity, has received relatively little attention.

Time-on-Task work described in literature revealed a research gap: Time-on-Task has not been studied in a systematic, comprehensive, and analytical way over the course of a school day. This paper introduced a novel ontological model for quantifying and analysing Time-on-Task losses in the school day at a level of such practical detail that it can be used to mobilise productivity improvement at the "Gemba", i.e. the place where the inputs of teaching are converted into the outputs of learning: The school classroom. Apart from its novel practicality, the model is also more systematic in its approach than the Time-on-Task studies to date (Stallings et al., 2014), due to its grounding in the well-established and powerful Overall Equipment Effectiveness (OEE) framework.

Developed through the Relevance-Design-Rigor Cycles (Hevner, 2007) of DSR, the model was initiated by the management of a large schools group (Relevance) , grounded in established theory (Rigor) and iteratively designed and refined through existing literature and the semantics of the target domain (Design). Once designed, the artifact was returned for testing in the target domain through focus groups of a total of 52 education stakeholders, of whom 39 were teachers and the rest were academics, managers, and specialists.

The result of the research was thus two-fold: the design artifact itself, and the results it achieved within the target domain. After being presented with the outputs of the TOTA model, 94.2% of the education stakeholders reported that the presentation had stimulated improvement thinking with them and 90.3% reported that they found that the TOTA provided a novel perspective to the school day. 80.8% of the sample said such studies would be beneficial in their school(s), with the teacher group surprisingly being slightly more enthusiastic than the rest of the stakeholder group to welcome further such studies.

As discussed in Table 1 of Chapter 3, the requirements for successful DSR were met, with only two exceptions: The need for further longitudinal studies using the TOTA model to test its effectiveness in translating created improvement energy into actual improvement, and the requirement to test the TOTA model output against those of competing models, which were not found in literature. The former requirement is addressed in section 6.2 below.

The significance of the research contribution lies in the addition of a useful construct to the relatively sparsely populated field of quantitative classroom observation studies with a time-loss focus (Apter et al., 2020, Wragg, 2011). Furthermore, the model's grounding in OEE makes it more systematic than previous Time-on-Task studies, and its schools level focus means that that it decentralises and energises improvement efforts at the Gemba: the school classroom. The research question was: *How can Time-on-Task in the school day be systematically, and comprehensively, analysed to enable teachers and school management to improve their school's productivity?*

The TOTA model, presented here for the first time in the public domain, presents a novel and valid tool, that was found to be novel, useful and beneficial by a stakeholder group notoriously resistant to change (Senge, 1995, Ehrenberg, 1999, Elmore, 2004).

The TOTA model was developed to measure and prioritise specific time losses within the school day at a level of detail relatable to teachers and principals, who are also in the best possible position to influence these losses. It is a productivity improvement enabler for the grass roots level. How it should be rolled-out at this level, is discussed next.

6.1 A Guideline for TOTA Model Implementation

During a TOTA study, data should be collected as a simple time-series observation data set, only noting a description of the activity observed and its start time, from which the duration of the activity can then later be derived by subtracting the start time from the start time of the next activity. It is recommended that the data be collected directly in a spreadsheet, using the time stamp function for the start times. After data collection, the activities and their durations can be coded into the different TOTA categories inserted into the columns of the spreadsheet, after which descriptive analytics can be done on the data using the TOTA Model structure and various levels of detail (Figure 3) to provide different Time-on-Task perspectives.

The allocation of the observations to various sub-categories within the TOTA model should be done by an analyst trained in the definitions of the TOTA categories, as described in Table 1. OEE scholars confirm that the complexity in accurate allocation of the data into is not to be underestimated (Hansen, 2002, Muchiri and Pintelon, 2008).

Due to this complexity, and in keeping with the people involvement principle found in the Lean Management theory, it is not recommended that the TOTA model be used as a top-down benchmarking tool, to compare the Time-on-Task ratios of different schools for example. However, in the case where it is used as such, the observations should be done by an objective party to the schools involved, trained in the TOTA model categories. If more than one such study officer is used, consistency in interpretation and method should be carefully managed by regularly calibrating interpretation among the officers, as well as with the theoretical definitions of the loss categories.

In contrast, however, when the TOTA model is used internally by teachers and principals as an analytical tool for improvement, the allocation of individual observations is less critical as the goal is to create awareness of, and prioritise, major loss categories present in the school day. In such a case, consistency between rounds of measurement (bi-annual or annual studies, for example) should be ensured.

Disseminating the analysis resulting from such studies is critical to the improvement effort. Whether teachers and principals are the study officers or not, they hold the key to reducing the Time-on-Task losses in their school, and therefore must be the target audience. The TOTA Model was not developed for high-level benchmarking studies, although it can be used as such, but rather to fill a gap in the field of Time-on-Task, namely measuring Time-on-Task in a way that enables improvement.

6.2 Further Research

The TOTA model has been tested in a pilot study situation and has shown to be able to energise teachers and other education stakeholders towards productivity improvement. Testing whether the energy created can be translated into actual productivity improvement, would be a source of further research. Repeated TOTA studies could show whether, which, and to what extent, Time-on-Task losses have been reduced. Researchers may then experiment with different approaches to facilitating the improvement in between studies to see which productivity improvement approach yields the best results. A possible research question could be: *How can the TOTA model be used to effect productivity improvement at schools' level?*

In terms of the theoretical aspect of the TOTA model, the discussion on the TOTA model development have highlighted that the performance and quality aspects of OEE have not fully been translated yet. For example, if a learner has already grasped a concept, but is exposed to unnecessary repetition of that concept, this would constitute an idling (performance) loss for that learner. Conversely, a situation where a learner did not get enough repetition and hence did not understand or retain a specific concept, would produce a learning quality loss. These are potentially important, but difficult, Time-on-Task losses to measure. Research questions to answer could be: *How can the TOTA model be expanded to include idling losses? How can the effectiveness of Time-on-Task in the school day be measured to expand the TOTA model to include a quality aspect?*

With Society 5.0 and Industry 4.0 unfolding, the education sector needs to play an agile, efficient, and effective upstream role to the workplace. There is no time to waste, so to speak. The TOTA model enables teachers and principals to identify, quantify and prioritise time losses within the school day - those activities that do not contribute to effective learning of value-adding skills and knowledge – in pursuit of building agile, efficient and effective schools.

AHLGRIM, R. W. 2010. *A thorough and efficient education: School funding, student achievement and productivity*, Indiana State University.

APTER, B., SULLA, F. & SWINSON, J. 2020. A review of recent large-scale systematic UK classroom observations, method and findings, utility and impact. *Educational Psychology in Practice,* 36**,** 367-385.

ASHTON, D. & GREEN, F. 1996. *Education, training and the global economy*, Edward Elgar Cheltenham.

BERGER, N. & FISHER, P. 2013. A well-educated workforce is key to state prosperity. *Economic Policy Institute,* 22**,** 1-14.

BOONE, H. N. & BOONE, D. A. 2012. Analyzing likert data. *Journal of extension,* 50**,** 1-5.

BROWN, I. & JAYAKODY, R. 2008. B2C e-commerce success: A test and validation of a revised conceptual model. *The Electronic Journal Information Systems Evaluation,* 11**,** 167-184.

BUTLEWSKI, M., DAHLKE, G., DRZEWIECKA-DAHLKE, M., GÓRNY, A. & PACHOLSKI, L. Implementation of TPM methodology in worker fatigue management-a macroergonomic approach. International Conference on Applied Human Factors and Ergonomics, 2017. Springer, 32-41.

CARROLL, J. 1963. A model of school learning. *Teachers college record,* 64**,** 723-723.

CREIGHTON, A. 2016. While Services Sector Booms, Productivity Gains Remain Elusive. *Wall Street Journal - Online Edition*.

DE RON, A. & ROODA, J. 2006. OEE and equipment effectiveness: an evaluation. *International Journal of Production Research,* 44**,** 4987-5003.

DELONE, W. H. & MCLEAN, E. R. 1992. Information systems success: The quest for the dependent variable. *Information systems research,* 3**,** 60-95.

DILLON, M., KOKKELENBERG, E. C. & CHRISTY, S. M. 2002. *The Effects of Class Size on Student Achievement in Higher Education Applying an Earnings Function,* [Place of publication not identified], Distributed by ERIC Clearinghouse.

EHRENBERG, R. G. 1999. Adam Smith goes to college: An economist becomes an academic administrator. *Journal of Economic Perspectives,* 13**,** 99-116.

ELMORE, R. F. 2004. *School reform from the inside out: Policy, practice, and performance*, ERIC.
FACCHINETTI, T. & CITTERIO, G. 2022. Application of the Overall Equipment Effectiveness to a Service Company. *IEEE Access,* 10**,** 106613-106640.
FISHER, C. W., FILBY, N., MARLIAVE, R., CAHEN, L., DISHAW, M., MOORE, J. & BERLINER, D. 1978. Beginning teacher evaluation study (Technical Report, V-1). *San Francisco, CA: Far West Laboratory*.
GERRITSEN, S., PLUG, E. & WEBBINK, D. 2017. Teacher Quality and Student Achievement: Evidence from a Sample of Dutch Twins. *Journal of Applied Econometrics,* 32**,** 643-660.
GETTINGER, M. & SEIBERT, J. K. 2002. Best practices in increasing academic learning time.
GIBBONS, P. M. & BURGESS, S. C. 2010. Introducing OEE as a measure of lean Six Sigma capability. *International Journal of Lean Six Sigma*.
HANSEN, R. C. 2002. Overall equipment effectiveness : a powerful production/maintenance tool for increased profits. 1st ed. ed. New York, NY: Industrial Press.
HEERINGA, S. G., WEST, B. T. & BERGLUND, P. A. 2017. *Applied survey data analysis*, chapman and hall/CRC.
HEVNER, A. R. 2007. A three cycle view of design science research. *Scandinavian journal of information systems,* 19**,** 4.
HEVNER, A. R., MARCH, S. T., PARK, J. & RAM, S. 2004. Design science in information systems research. *MIS quarterly***,** 75-105.
HOLLINGSWORTH, J. & YBARRA, S. 2006. Methods for generating classroom productivity index. Google Patents.
HOXBY, C. M. 2003. School choice and school productivity. Could school choice be a tide that lifts all boats? *The economics of school choice.* University of Chicago Press.
HOXBY, C. M. 2004. Productivity in Education: The Quintessential Upstream Industry. *Southern Economic Journal,* 71**,** 209-231.
IVANCIC, I. Development of maintenance in modern production. Euromaintenance'98 Conference Proceedings, 1998. CRO Dubrovnik/Hrvatska, 5-7.
JANSEN, I. P., RAMNATH, S. & YOHN, T. L. 2012. A diagnostic for earnings management using changes in asset turnover and profit margin. *Contemporary Accounting Research,* 29**,** 221-251.
KOBAYASHI, I. 1995. *20 keys to workplace improvement,* Portland, Ore, Productivity Press.
KRAFT, M. A. & MONTI-NUSSBAUM, M. 2020. The Big Problem with Little Interruptions to Classroom Learning. *Annenberg Institute at Brown University*.
LAFORTUNE, J., ROTHSTEIN, J. & SCHANZENBACH, D. W. 2018. School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics,* 10**,** 1-26.
LITTLE, S. G. & AKIN-LITTLE, A. 2008. Psychology's contributions to classroom management. *Psychology in the Schools,* 45**,** 227-234.
MERRIAM-WEBSTER, D. 2002. Merriam-webster. *On-line at http://www. mw. com/home. htm.*
MUCHIRI, P. & PINTELON, L. 2008. Performance measurement using overall equipment effectiveness (OEE): literature review and practical application discussion. *International journal of production research,* 46**,** 3517-3535.
MUÑOZ-VILLAMIZAR, A., SANTOS, J., MONTOYA-TORRES, J. R. & JACA, C. 2018. Using OEE to evaluate the effectiveness of urban freight transportation systems: A case study. *International Journal of Production Economics,* 197**,** 232-242.
PRATER, M. A. 1992. Increasing time-on-task in the classroom. *Intervention in School and Clinic,* 28**,** 22-27.
RAOUF, A. 1994. Improving capital productivity through maintenance. *International Journal of Operations & Production Management*.

SANDERS, W. L., WRIGHT, S. P. & HORN, S. P. 1997. Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation. *Journal of Personnel Evaluation in Education,* 11**,** 57-67.

SARASON, S. B. 1990. *The Predictable Failure of Educational Reform: Can We Change Course before It's Too Late? The Jossey-Bass Education Series and the Jossey-Bass Social and Behavioral Science Series*, ERIC.

SCHREYER, P. & PILAT, D. 2001. Measuring productivity. *OECD Economic studies,* 33**,** 127-170.

SENGE, P. 1995. On schools as learning organizations: A conversation with Peter Senge. *Educational Leadership,* 52**,** 20-23.

SHANKS, G., TANSLEY, E. & WEBER, R. 2003. Using ontology to validate conceptual models. *Communications of the ACM,* 46**,** 85-89.

SHINGO, S. 1996. *Quick changeover for operators: the SMED system*, Productivity press.

SIMON, H. A. 1996. *Models of my life*, MIT press.

SPÖTTL, G. & WINDELBAND, L. 2021. The 4th industrial revolution–its impact on vocational skills. *Journal of Education and Work,* 34**,** 29-52.

STALLINGS, J. 1980. Allocated academic learning time revisited, or beyond time on task. *Educational researcher,* 9**,** 11-16.

STALLINGS, J., ALMY, M., RESNICK, L. B. & LEINHARDT, G. 1975. Implementation and child effects of teaching practices in Follow Through classrooms. *Monographs of the society for research in child development***,** 1-133.

STALLINGS, J., NEEDELS, M. & STAYROOK, N. 1979. The teaching of basic reading skills in secondary schools, Phase II and Phase III. *Menlo Park, CA: Stanford Research Institute*.

STALLINGS, J. A., KNIGHT, S. L. & MARKHAM, D. 2014. Using the stallings observation system to investigate time on task in four countries.

VENABLE, J. R. Design science research post Hevner et al.: Criteria, standards, guidelines, and expectations. International Conference on Design Science Research in Information Systems, 2010. Springer, 109-123.

WALBERG, H. J. 1984. Improving the productivity of America's schools. *Educational leadership,* 41**,** 19-27.

WRAGG, T. 2011. *An introduction to classroom observation (Classic edition)*, Routledge.