

Measuring English teaching quality: what teacher effectiveness research can offer?

Siti Nurul Azkiyah – Amirul Mukminin

DOI: 10.18355/XL.2023.16.02.18

Abstract

Research findings worldwide suggest that teachers should continually improve their teaching quality, which refers to instruction or actions to deliver teaching and learning materials, even if they are good teachers. The question is the basis that should be considered to decide the improvement priorities. Therefore, this paper reviews theoretically driven and empirically validated teacher instruction contributing to student outcomes, which can serve as a framework for deciding the areas of improvement. The paper also reviews relevant techniques to measure teacher instruction to determine teachers' teaching quality. The study employed a qualitative method, in which a literature review on educational effectiveness research, especially teacher effectiveness research, was carried out to understand teacher actions in the classroom that contribute to achieving learning goals, especially concerning English lessons, as the focus of this paper. The paper used keywords such as effective English teaching, educational effectiveness research, teaching quality, teacher effectiveness research, and techniques of measuring teaching quality to find relevant literature. The findings conclude that the classroom factors of the dynamic model of educational effectiveness research could serve as a framework to conceptualize English teaching quality, whereas techniques like student surveys or observation can be used to investigate teachers' initial instructional practices.

Key words: educational effectiveness research, effective English teaching, teaching quality, teacher effectiveness research, techniques to measure teaching quality

Introduction

Every teacher needs to improve their quality, including good teachers since teaching quality can fluctuate, and good teachers must maintain their quality. Moreover, research across different countries (e.g., Azigwe, 2016; Azkiyah, & Mukminin, 2017; Ehren, 2015; Goldhaber, 2015; Chuyun Hu, 2020; Mukminin et al., 2017; Reynolds et al., 2014; Scheerens, 2013; Stronge, 2018; van de Grift, 2014) have found that teachers are the most critical actors in education. Hence, it is deemed significant to continually assess teacher quality, including English teachers, since the assessment is crucial not only to understand their competencies but also to serve as a baseline to decide on priority areas of teacher improvement. The question is what framework should be used to fairly and comprehensively assess teacher quality. In this case, it is relevant to clarify the components of teacher quality, especially those empirically proven to be positively related to better student learning outcomes, especially in the English subject. Therefore, this study offers a literature review where several keywords, including teacher quality, teacher effectiveness research, English teaching, and methods of measuring teacher quality, are used to search relevant literature.

It is first vital to note that the concept of teacher quality has been recognized as a broad term; hence it takes work to define. The term teacher quality is often associated with effective teaching, which according to Chuyun Hu (2020), is a decade-old question among educational researchers and teachers. Moreover, the concept significantly varies among students and teachers (Miron & Mevorach, 2014) and therefore is hard to define. However, concerning the components of teacher quality, some studies in different countries, such as the United States (Goldhaber, 2015; Darling-Hammond, 2010; Jacob, 2012; Nava et al., 2019), Latin America (Hunt,

2009), Europe (van de Grift, 2014; Creemers & Kyriakides, 2008), South Korea (van de Grift, Chun, Maulana, Lee, & Helms-Lorenz, 2017) as well as Indonesia (Ree & Jaitse, 2016) showed various aspects such as teacher level of education and years of teaching experience, the mastery of subject knowledge, teaching or pedagogical knowledge, teacher certification, and teacher instruction in the classroom.

Among these aspects, van de Grift (2014) reviews some previous studies documenting that around 15 – 25% of differences in student learning outcome is explained by some teaching elements such as curriculum, creating a safe and stimulating environment, efficient classroom management, and teaching-learning strategies, all of which can be referred to as teacher instruction. Moreover, they are characteristics that can be interfered with through intervention programs through which a connection between research or theories and improvement practices can be encouraged. Therefore, the most critical criterion in conceptualizing teacher quality is teacher instruction in the classroom.

Moreover, the findings of previous studies on the effects of other factors, such as the teacher's level of education and years of teaching experience, yielded different results. For instance, old yet famous reviews conducted by Hanushek (1989, 1997) found no strong evidence of the effects of teacher education and years of teaching experience on student outcomes. However, other studies showed a strong relationship between teacher education, years of teaching experience, and student achievement (Greenwald, Hedges, & Lain, 1996; Hedge, Lain, & Greenwald, 1994). A more recent review by Jacob (2012) concluded that the first few years of teaching experience positively relate to effective learning outcomes, yet the benefits of years of teaching experience plateau quickly. In addition, inconclusive results were reported by Dial (2008), where teacher degree level alone did not affect student achievement and had positive effects when combined with years of teaching experience.

Other two aspects studied concerning teacher quality are teacher subject knowledge and pedagogical knowledge, which previous findings also show different results, not to mention contradictory. While some studies suggested a positive relationship between teacher subject knowledge and student achievement (e.g., Darling-Hammond, 2000; Goldhaber & Brewer, 2000; Tchoshanov, 2011), other studies did not show such a relationship (Rowan, Correnti, & Miller, 2002; Monk, 1994). Furthermore, reviewing various studies, Fung et al. (2017) concluded that while subject-matter knowledge has been generally linked to student performance in some subjects (mainly Math), the relationship has often been weak or statistically insignificant. Instead, according to their study, high levels of pedagogic efficacy are the ones that positively affect student achievement. Concerning teacher certification, previous studies confirm mixed evidence that teachers passing the certification assessment are more effective (e.g. Cowan & Goldhaber, 2016). Furthermore, Goldhaber (2015) reported that the relationship between teachers' certification status, especially in the case of the US, is bound to vary depending on the state, subject area, and grade level due to the differences in certification requirements along these dimensions.

The above literature review shows that teacher level of education, subject and pedagogical knowledge, and years of teaching experience and certification are not consistently found to be positively related to student outcomes. Therefore, it is crucial to consider other factors which have consistently contributed to better student outcomes. In this case, as previously mentioned, studies across different countries have revealed the significant role of teacher instruction in the classroom. Therefore, this paper employs teacher instructions as the primary indicator of teacher quality because, in this way, teacher quality has more grounded evidence and can be intervened through improvement processes.

Concerning teacher instructions in the classroom, the findings of Educational Effectiveness Research, especially Teacher Effectiveness Research (TER), which concerns teacher actions that positively contribute to student outcomes, are worth considering when conceptualizing teacher quality. Teacher actions, in this case, are considered to be teacher instruction, which refers to teachers' efforts in maximizing students' learning opportunities. Hence, teacher instructions and actions are used interchangeably in this paper. This paper will first address teacher instructions based on TER, which attempts to provide a concept of teacher quality and simultaneously answer the question of what to assess when assessing teacher quality. Then, the paper recommends the strategies or techniques for measuring teacher quality and its instrument, which answers the question of how to assess teacher quality. Finally, it is essential to note that English is the focus of the study, and therefore a more detailed explanation of effective English instructional strategies is discussed.

What to Assess

The Findings of Teacher Effectiveness Research (TER)

The findings of TER have shown a number of teacher actions in the classroom related to student outcomes. Therefore they can be used as a framework for deciding teacher actions that should be observed when measuring teaching quality. At this point, it is essential to clarify that student outcomes have been largely using student cognitive outcomes as an indicator, although some experts (e.g., Creemers & Kyriakides, 2008; Van der Werf, Opdenakker, & Kuyper, 2008) have criticized this to narrow down the scope and the meaning of education. Nevertheless, other studies have shown the small impact of education on the non-cognitive outcome, possibly due to the more significant effects of outside school factors on student non-cognitive outcomes (van der Werf et al., 2008). Thus, despite being considered a traditional approach, using student cognitive outcome as the indicator of effective teacher instruction is essential and reasonable.

Reviewing studies conducted during the 1970s, 1980s, and 1990s, Muijs and Reynolds (2011) presented a review of teacher instructions found to positively affect student outcomes, which amount to 60 characteristics, such as the establishment of effective school and classroom climate, the creation of students' learning opportunities and classroom management and effective questioning, feedback, and reinforcement. In creating learning opportunities, several activities have been identified and empirically validated, such as raising high expectations (Marzano, 2000), emphasis on academic goals and achievement (Cotton, 1995; Doyle, 1986; Powell, 1980), clear and step-wise presentation of materials as well as effective questioning and feedback (Brophy & Good, 1986; Kane et al., 2010; Scheerens & Bosker, 1997; Seidel & Shavelson, 2007), and clear structures and routines (Brophy & Good, 1986).

Effective teaching does not happen when teachers explain the materials and students take notes; instead, effective teaching occurs when teachers allow students to share opinions and feedback and engage in the teaching and learning process. In other words, a discussion is a preferred technique of teaching effective teachers use. In addition, effective teachers respect students' various learning styles and opinions (Cook-Sather et al., 2015; Miron & Mevorach, 2014), and therefore they offer a safe and respectful learning environment with plenty of chances for students to share their knowledge and express their opinions (Cook-Sather et al., 2015; Scarboro, 2012; Stronge, 2018).

Next, in line with the need to respect individual students and considering the demand for teachers to adjust their instructions according to students' needs, differentiated or adaptive instruction has been positively found to contribute to student achievement (Creemers & Kyriakides, 2008; Deunk et al., 2018; Steenbergen-Hu, Makel, &

Olszewski-Kubilius (2016). Differentiated instruction can happen both within and beyond the classroom, although the ones implemented in the classroom receive more attention. Within the classroom, Maulana, Sale-Jacobse, & Helms-Lorenz (2019) suggest that teachers can adjust the complexity of the tasks, offer different types of feedback and learning time, and allow variation in outputs for different groups of students. Only some teachers can perform this complex teaching skill, and it is challenging to implement. It is part of highly demanding teaching behaviors commonly implemented only by highly effective teachers (Pietsch, 2010; van der Lans et al., 2017). In other words, only teachers with good teaching quality are more likely to incorporate differentiated instruction into their teaching practice (Pietsch, 2010).

Furthermore, to contribute to the development of effectiveness theories, Creemers and Kyriakides (2008) developed the dynamic model of educational effectiveness research, which is a further development of effective classrooms established by Creemers (1994). The model has four levels, namely context/policy level, school level, teacher level, and student level (See Figure 1). Referring to the superiority of teachers as found by TER, the dynamic model emphasizes the importance of the classroom level and expects the higher level, i.e., the context/policy and school levels, to provide the necessary conditions for the effectiveness of the classroom level

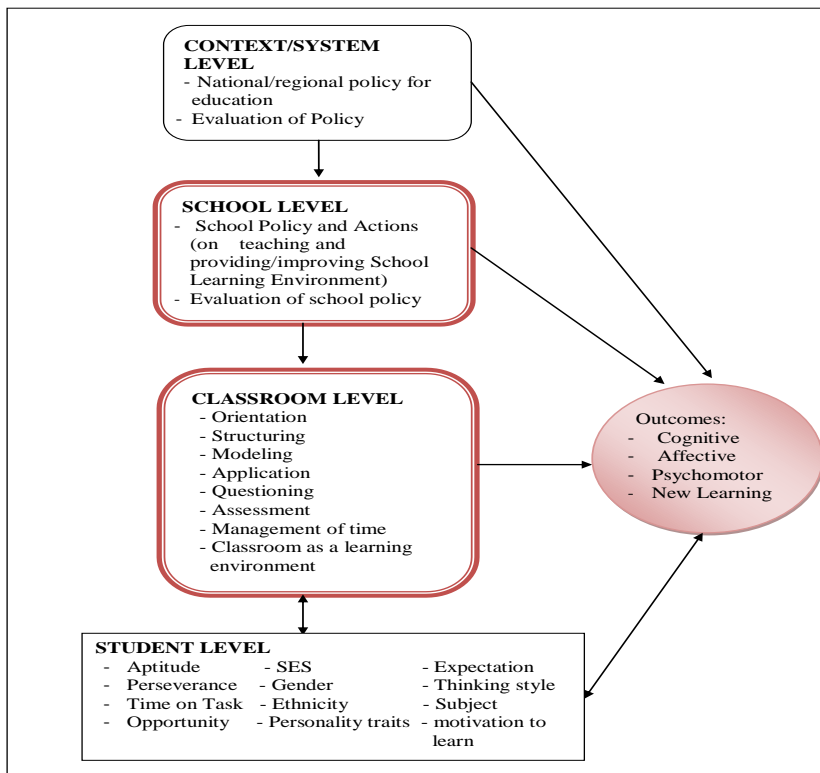


Figure 1 The dynamic Model of Educational Effectiveness Research (2008)

Referring to the findings of TER, eight factors are included in the classroom level of the dynamic model, 1) orientation, 2) structuring, 3) questioning, 4) teaching modeling, 5) application, 6) management of time, 7) the classroom as a learning environment (CLE), and 8) classroom assessment. Orientation is the teacher's attempt to explain the importance of the lesson to students' daily life, which is expected to help the students understand the importance of their learning activities. Structuring concerns teachers' explanation of the series of activities during the lesson and teachers' attempt to arrange the lesson that proceeds from easier to more difficult ones. With regard to English subjects, especially for reading comprehension, structured instruction has been found to be related to student achievement (Kane, Taylor, Tyler, & Wooten, 2010).

Furthermore, questioning refers to teachers' attempts to provide varieties of questions in terms of difficulty level (easy and difficult) and type (process and product) and appropriate reactions to students' responses. Modeling is the provision of learning strategies or the encouragement of students to develop their own so that self-directed learning can be promoted. The application relates to the immediate exercise of the topics taught during the lesson. Management of time concerns teachers' attempts to organize their lessons in such a way as to ensure that students are on task throughout the lesson and that they are engaged during the lesson. CLE includes the following components: 1) teacher-student interaction, 2) student-student interaction, 3) students' treatment by the teachers, 4) competition among the students, and 5) classroom disorder. Finally, assessment deals with teachers' attempts to collect information on their students' knowledge and skills to identify their learning needs.

The above classroom factors of the dynamic model are all teacher instructions that have been found to be positively related to student outcomes and are in line with other previous studies. Therefore, this paper proposes the use of the classroom factors of the dynamic model to serve as the basis to indicate teacher quality since the model offers a comprehensive model of effectiveness. Several studies have been reported to validate the model through various research designs, one of which is experimental design (Antoniou, 2009). In addition, the model requires not only the knowledge of the subject matter but also that of pedagogy, which has been used as an indicator of teacher quality. Teachers can only provide modeling and raise various questions if they have sufficient knowledge of the subject matter. In the context of English teaching, Schachter & Freeman (2020) have argued that the English fluency of English teachers matters a lot.

Similarly, teachers will only be able to manage the classroom to enable students to have maximum learning opportunities and to engage them during the learning process with enough pedagogical knowledge. Since this paper focuses on English subjects, the following will explain the use of TER, especially the classroom factor of the dynamic model as a framework in formulating teacher instructions that should be taken into account when measuring teacher quality.

The Use of TER as a Framework in Measuring English Teaching Quality

In measuring English teaching quality, several activities that English teachers may do in the classroom according to the eight classroom factors of the dynamic model should be first identified, the main descriptors of which are provided in Table 1. This can serve as a framework for deciding what to investigate when measuring English teaching quality. Nevertheless, it is crucial to note that some of them are generic and can therefore be used for other subjects as well.

Table 1 The Classroom Factors and Their Descriptors for English Subject (adapted from Creemers & Kyriakides, 2008)

Factors	Descriptors
1) Orientation	a) Whether teachers explain the importance of the tasks/topics / activities of the lesson by connecting the lesson to students' daily life, previous lessons, or other subjects; b) Whether teachers challenge or invite students to identify the reasons why they learn certain topics such as transportation, weather, descriptive text, and narrative text.
2) Structuring	a) Whether teachers starts the lesson with overviews and/or review of objectives; b) Whether teachers outline the contents to be covered and signal transitions between lesson parts; c) Whether teachers call attention and review main ideas.
3) Teaching modelling	a) Whether teachers use specific strategies of teaching listening, reading, writing and speaking; b) Whether teachers invite students to demonstrate strategies of learning they may know.
4) Questioning	a) Whether teachers raise different types of questions (i.e. process and product) at appropriate difficulty level; b) Whether teachers give enough time to students to respond; c) Whether teachers dealing with student responses appropriately, such as giving clue when students cannot answer questions.
5) Application	a) Whether teachers provide appropriate and sufficient tasks / exercises; b) Whether teachers use seatwork or small group tasks in order to provide needed practice and application opportunities; b) Whether teachers develop application tasks replicating what have been presented in modeling or more challenging as starting points for the next step of teaching and learning.
6) The classroom as a learning environment	a) Whether teachers establish on task behavior through the interactions they promote (i.e., teacher-student and student-student interactions); b) Whether teachers deal with classroom disorder and student competition through establishing rules, persuading students to respect them and using the rules.
7) Management of Time	a) Whether teachers organize the classroom environment and maximize engagement rates of students.
8) Assessment	a) Whether teachers use appropriate techniques to collect data on student knowledge and skills; b) Whether teachers analyze data to identify student needs and report the results to students and parents; c) Whether teachers use the results of assessment to evaluate their own practices.

In the opening session, it is important to look at orientation activities, where teachers are expected to explain the importance of the lesson by connecting the topic to students' daily life or previous topics. For instance, when the topic is related to various means of transportation, teachers can ask how students come to school, expecting that students will mention different means of transportation. It is also possible that teachers show some pictures showing the varieties of vehicles students often see in their environment. In this way, teachers facilitate students to know some vocabulary they will learn and, at the same time to be aware of the importance of the topic in their daily life. It is also possible to explain the importance of the lesson by challenging students to identify the advantages of studying means of transportation.

In the next stage, students should be aware of the goals and the activities prepared by their teachers to accomplish the goals. Therefore, it is crucial to investigate whether teachers follow orientation activities by presenting the goals of the lesson and the activities intended to reach the goals. In this way, students can engage in monitoring the stages of the lesson through which students learn to regulate their own learning. Although this activity seems to be simple, the study by Azkiyah (2017) found that some English teachers forgot to explain the goals and the activities during the lesson. Instead, they commonly ask students to open their books, and therefore it is suggested that English teachers explain the goals and the activities of the lesson as well as signal the transition from one activity to the next.

Furthermore, in delivering the materials, effective English teachers are expected to use various modeling activities to teach different skills or language components to be taught. The use of modeling activities can promote students' autonomous learning since they know the strategies to accomplish exercises. In teaching reading, for instance, pre-reading activities can be done by drawing a semantic map to brainstorm what students have known and need to know about the topic. During reading activities, can use different strategies depending on the sub-reading skills teachers decide as the goals of the lesson. The strategies for teaching how to find main ideas and answer WH (where, when, what, who) types of questions can be different. Therefore, at this stage, it is relevant to analyze whether teachers apply appropriate teaching strategies.

Then, to guide the discussion in the classroom, English teachers can initiate both product and process questions, referred to as questioning. In reading comprehension, product questions are questions whose answers can be explicitly found in the text, whereas process questions are those requiring students to think beyond the printed information in the text and hence promote students' critical thinking. In order to deepen students' understanding, quality English teachers provide relevant exercises to reach the goals of the lesson, which the dynamic model used in this paper refers to as an application. In developing application tasks, it is useful to consider whether the tasks replicate what has been previously explained during modeling activity or more complex ones.

The next characteristic of effective teachers is monitoring whether students are on task throughout the lesson, promoting collaboration and competition, can manage classroom disorder, and maximizing student engagement. The dynamic model considers these activities as the establishment of the classroom as a learning environment and time management. In this case, the measurement of English teaching quality should focus on the aforementioned activities to ensure that learning opportunities are provided and students can benefit from the process. Finally, assessment, especially the formative one, is an important characteristic of an effective teacher that English teachers should do throughout different stages of the lesson. The activities can vary, including reviewing the materials before closing, asking students to reflect or write a summary, distributing worksheets, posing open-ended questions, or using self and peer assessment.

When English teachers have done those activities, it is safe to conclude that they are quality English teachers. The areas where teachers still need to implement or are considered to be lacking become the priorities of improvement designed through a teacher development program. In this way, teacher development can be designed according to the current teachers' competence and, at the same time, directed toward theoretically and empirically proven practices of effective teachers. Thus, the first concern of this paper has been addressed, while the remaining question is how to measure teachers' teaching quality, which is presented in the following section.

How to Assess

This section discusses techniques or strategies and instruments that should be used to collect data on teacher instructions indicating their quality. Previous studies have used various techniques, such as videotaped observation, student surveys, and value-added models (e.g., Goldhaber, 2015). Compared to student surveys, the use of observation by trained observers is likely to give more independence in rating teachers' teaching quality since some cultures encourage students to respect their teachers. It will be considered impolite when students show their teachers' weaknesses in their teaching practice. In the Indonesian context, for instance, as revealed by Maulana, Opendakker, Den Brok, and Bosker (2010), the teaching profession is considered to be a high and respectful occupation, and therefore teachers benefit from being respected by both

students and society in general. The study of Azkiyah (2013) showed that students rate their teachers better compared to independent observers.

However, the use of classroom observation by trained observers is challenging. Rater reliability has been widely acknowledged to be a persistent problem, and underlying factors include variation in severity and the changes in raters' severity over time (Casabianca, Lockwood, & McCaffrey, 2015). In addition, a series of training needs to be conducted to facilitate common understanding among observers. Both the training and the implementation of classroom observation are, of course, costly. On the other hand, the use of a student survey is relatively easy because training is not required and can involve the whole number of students in the class.

Next, the value-added model concerns the use of student achievement as an indicator of teachers' teaching quality, in which better student achievement is the indicator of better teachers' teaching quality. In addition, the self-reported measure is another way of evaluating teaching quality, in which teachers have to fill in a questionnaire to indicate the quality of their own teaching practices. The study of Irwin, Madura, Bamat, and McDermott (2016), however, showed that classroom observation measures distinguish teachers' teaching quality better than self-reported measures do. This is logical since, in most cases, teachers, especially the experienced ones, tend to consider themselves better teachers. Furthermore, given the broad spectrum of teaching quality, there are, of course, other techniques for measuring teaching quality, such as principal evaluation, analysis of classroom artifacts, and interviews.

Considering the above literature and following the instrument suggested by the dynamic model of educational effectiveness research, this paper proposes the use of classroom observation and student questionnaires as techniques for measuring teachers' teaching quality. Classroom observation conducted by trained independent observers enables the gathering of significant and useful information about teachers' practice when it is conducted carefully, and its data are used thoughtfully (Little, Goe, & Bell, 2009). In line with this, classroom observation also provides necessary information that other instruments cannot show, such as 1) strengths and shortcomings that teachers have, 2) baseline information to determine appropriate assistance and support to struggling teachers, and 3) serve as a formative tool for coaching teacher performance (Gallagher, Rabinowitz, & Yeagley, 2011; Goe & Croft, 2009). Furthermore, Gallagher et al. (2011) suggested using a student questionnaire because it is efficient both in terms of time and cost, easy to administer, and may build stronger positive relationships between teachers and students. Moreover, they reveal that student survey has been shown to be more strongly correlated with student achievement than an administrator or self-reported measures of teacher effectiveness.

Concerning the specific subject involved, including English, the most important aspect to be carried out is identifying possible activities that effective English teachers do in the classroom according to each classroom factor of the dynamic model. Starting from the orientation stage, when the topic is about a season, for instance, effective English teachers are expected to connect such topics to students' daily lives by asking students about the season in their surroundings and activities they may do in each season. In the modeling phase, effective English teachers are likely to use specific strategies for teaching various language skills and components or invite students to demonstrate strategies of learning they may know and use. In the instruments, the question of whether teachers connect the lesson to students' daily lives could be raised. In terms of modeling, a relevant question could be whether teachers explain some specific strategies that students can use in order to answer exercises usually given during application. In principles, all activities that effective teachers do should be first identified and then made in statements or listed in the

instruments. In this case, Table 1 is useful for considering different activities that effective English teachers do in their classrooms.

Conclusion

This paper has addressed two important issues in assessing teachers' quality, namely, teacher performance that must be taken into account and the strategies to collect the data on teacher performance. The literature review suggests the use of the classroom factors of the dynamic model of educational effectiveness research as the framework for measuring teacher performance in the classroom. In addition, the review also proposes classroom observation and student surveys as strategies to assess teacher performance due to the benefits of both techniques. Thus, the findings of this study have provided evidence-based guidelines for assessing teacher quality, both in terms of the aspects and the strategies. Finally, the data gathered during the assessment of teacher quality could serve as a baseline study, which informs teachers' strengths and weaknesses and therefore provide essential information on relevant areas of improvement for teachers.

Bibliographic references

- Antoniou, P. (2009). Using the dynamic model of educational effectiveness to improve teaching practice: building an evaluation model to test the impact of teacher professional development programs (Unpublished PhD thesis). University of Cyprus: Department of Education.
- Azigwe, J. B. (2016). Using comparative international studies for modeling educational effectiveness: A secondary analysis of PISA-2009 study. *Journal of Education and Practice*, 7 (18), 199-209.
- Azkiyah, S. N. (2017). The Development of Effective English Teacher Improvement Program Based on a Case Study of English Teaching in Madrasah in Indonesia. *International Journal of Applied Linguistics & English Literature*, (6) 2, 1-7.
- Azkiyah, S. N. (2013). The effects of two intervention programs on teaching quality and student outcome: A comparison of education standards and education standards combined with a teacher improvement program in Indonesia (Unpublished Doctoral Dissertation). The University of Groningen: Groningen Institute for Educational Research (GION).
- Azkiyah, S. N., Doolaard, S., Creemers, B. P. M., & (Greetje) Van Der Werf, M. P. C. (2014). The effects of two intervention programs on teaching quality and student achievement. *Journal of Classroom Interaction*, 49 (1), 4-11.
- Azkiyah, S. N., & Mukminin, A. (2017). In search of teaching quality of EFL student teachers through teaching practicum: Lessons from a teacher education program. [V iskanju kakovosti poučevanja prihodnjih učiteljev angleščine kot tujega jezika med pedagoško prakso: Program izobraževanja učiteljev] *Center for Educational Policy Studies Journal*, 7 (4), 105-124.
- Brophy, J. E., & Good, T. L. (1986). Teacher behaviour and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching*, 328-375. New York: Macmillan.
- Casabianca, J. M., Lockwood, J. R., & McCaffrey, D. F. (2015). Trends in classroom observation scores. *Educational and Psychological Measurement*, (75) 2, 311-337.
- Chuyun Hu, C. (2020). Understanding college students' perceptions of effective teaching. *International Journal of Teaching and Learning in Higher Education*, 32 (2), 318-328. <https://eric.ed.gov/?id=EJ1286303>.
- Cook-Sather, A., Clarke, B., Condon, D., Cushman, K., Demetriou, H., & Easton, L. (2015). *Learning from the student's perspective: A sourcebook for effective teaching*. Abingdon, UK: Routledge.
- Cotton, K. (1995). *Effective student practices: A research synthesis*. Retrieved from <http://www.nwrel.org/scpd/esp/esp95.html#1>.

- Cowan, J., Goldhaber, D. (2016). National Board Certification and teacher effectiveness: Evidence from Washington State. *Journal of Research on Educational Effectiveness* (9) 3, 233-258.
- Creemers, B.P.M. (1994). *The effective classroom*. London: Cassel.
- Creemers, B.P.M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness*. London: Routledge.
- Darling-Hammond, L. (2000). Reforming teacher preparation and licensing: Continuing the debate. *Teachers College Record*, 102 (1), 5-27.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Archives*, 8 (1). Retrieved from <http://epaa.asu.edu/epaa/v8n1>.
- Darling-Hammond, L. (2010). *Evaluating teacher effectiveness: How teacher performance assessments can measure and improve teaching*. Washington, DC: Center for American Progress.
- De Ree, J. J. (2016). Indonesia - Teacher certification and beyond: An empirical evaluation of the teacher certification program and education quality improvements in Indonesia. Washington, D.C. : World Bank Group. <http://documents.worldbank.org/curated/en/129551468196175672/Indonesia-Teacher-certification-and-beyond-an-empirical-evaluation-of-the-teacher-certification-program-and-education-quality-improvements-in-Indonesia>
- Dial, J. C. (2008). *The effect of teacher experience and teacher degree levels on student achievement in Mathematics and Communication Arts* (Unpublished Dissertation). Baker University.
- Deunk, M. I., Smale-Jacobse, A. E., de Boer, H., Doolaard, S., & Bosker, R. J. (2018). Effective differentiation practices: a systematic review and meta-analysis of studies on the cognitive effects of differentiation practices in primary education. *Educational Research Review* (24), 31–54. <https://doi.org/10.1016/j.edurev.2018.02.002>.
- Doyle, W. (1986). *Classroom organization and management*. M. C. Wittrock (Ed.), *Handbook of research on teaching*, 392-431. New York: Macmillan.
- Ehren, J. S. M. (2015). The Evidence base for school inspection frameworks. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal)* 12-December 2015, 12, 43.
- Fung, D., Kutnick, P., Mok, I., Leung, F., Lee, B. P., Mai, Y. M., & Tyler, M. T. (2017). Relationships between teachers' background, their subject knowledge and pedagogic efficacy, and pupil achievement in primary school mathematics in Hong Kong: An indicative study. *International Journal of Educational Research* (81), 119-130.
- Gallagher, C., Rabinowitz, S., and Yeagley, P. (2011). *Key considerations when measuring teacher effectiveness: A framework for validating teachers' professional practices* (AACC Report). San Francisco and Los Angeles, CA: Assessment and Accountability Comprehensive Center.
- Goe, L., & Croft, A. (2009). *Methods of evaluating teacher effectiveness*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from http://tqcenter.learningpt.org/publications/RestoPractice_EvaluatingTeacherEffectiveness.pdf
- Goldhaber, D. (2015). *Teacher effectiveness research and the evolution of US teacher policy*. The Productivity for Results Series No. 5. George W. Bush Institute, Education Reform Initiative.
- Goldhaber, D., & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, 22 (2), 129–145.

- Greenwald, R., Hedges, L. V., & Lain, R. D. (1996). The effects of school resources on student achievement. *Review of Educational Research*, 66 (3), 361-396.
- Hanushek, E. A. (1989). The impact of differential expenditures on school performance. *Educational Researcher*, 18 (4), 45-65.
- Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, 19 (2), 141-164.
- Hedge, L. V., Lain, R. D., & Greenwald, R. (1994). Does money matter? A metaanalysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher*, 23 (3), 5-14.
- Hunt, B.C. (2009). *Teacher effectiveness: A review of the international literature and its relevance for improving education in Latin America*. Washington, DC: PREAL.
- Irwin, C. W., Madura, J. P., Bamat, D., & McDermott, P.A. (2016). Patterns of classroom quality in Head Start and Center-Based Early Childhood Education Programs. *Regional Educational Laboratory Northeast & Islands*.
- Jacob, A. (2012). Examining the relationship between student achievement and observable teacher characteristics: Implications for school leaders. *International Journal of Educational Leadership Preparation*, 7 (3), 1 – 13.
- Kane, T.J., Taylor, E.S., Tyler, J.H., & Wooten, A.L. (2010). Identifying effective classroom practices using student achievement data. *The Journal of Human Resources* 46 (3), 587 – 613.
- Little, O., Goe, L., & Bell, C. (2009). *A practical guide to evaluating teacher effectiveness*. Washington, DC: National Comprehensive Center of Teacher Quality.
- Marzano, R. J. (2000). *Transforming classroom grading*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Marzano, R. J. (2007). *The art and science of teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Maulana, R., Opdenakker, M., den Brok, P., & Bosker, R. (2011). Teacher-student interpersonal relationships in Indonesia: Profiles and importance to student motivation. *Asia Pacific Journal of Education*, 31 (1), 33-49.
- Miron, M., & Mevorach, M. (2014). The "good professor" as perceived by experienced teachers who are graduate students. *Journal of Education and Training Studies*, 2 (3), 82-87.
- Monk, D. H. (1994). Subject area preparation of secondary mathematics and science: Teachers and student achievement. *Economics of Education Review*, 13, 125-145.
- Muijs, D., & Reynolds, D. (2011). *Effective teaching: Evidence and practice* (3rd ed.). London: SAGE.
- Mukminin, A., Kamil, D., Muazza, M., & Haryanto, E. (2017). Why teacher education? documenting undocumented female student teachers' motives in indonesia: A case study. *Qualitative Report*, 22 (1), 309-326.
- Nava, I., Park, J., Dockterman, D., Kawasaki, J., Schweig, J., Quartz, K. H., & Martinez, J. F. (2019). Measuring Teaching Quality of Secondary Mathematics and Science Residents: A Classroom Observation Framework. *Journal of Teacher Education*, 70(2), 139-154. <https://doi.org/10.1177/0022487118755699>
- Powell, A. (1980). *The uncertain profession: Harvard and the search for educational authority*. Cambridge, Massachusetts: Harvard University Press.
- Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C., & Stringfield, S. (2014). Educational Effectiveness Research (EER): A state-of-the-art review. *School Effectiveness and School Improvement*, 25 (2), 197-230.
- Rowan, B., Correnti, R., & Miller, R. (2002). What large scale survey research tells us about teacher effects on student achievement: Insights from the prospects study of elementary schools. *Teachers College Record*, 104 (8), 1525–1567.
- Scarboro, A. (2012). Student perception of good teaching. *International Journal of New Trends in Arts, Sports & Science Education (IJTASE)*, 1 (1), 49-66.

- Scheerens, J. (2013). The use of theory in school effectiveness research revisited. *School effectiveness and school improvement*, 24 (1), 1-38.
- Scheerens, J., & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77 (4), 454-499.
- Steenbergen-Hu, S., Makel, M. C., & Olszewski-Kubilius, P. (2016). What one hundred years of research says about the effects of ability grouping and acceleration on K–12 students' academic achievement: findings of two second-order meta-analyses. *Review of Educational Research*, 86 (4), 849-899. <https://doi.org/10.3102/0034654316675417>.
- Stronge, J. H. (2018). *Qualities of effective teachers*. Alexandria, VA: ASCD.
- Tchoshavov, M. A. (2011). Relationship between teacher knowledge of concept and connections, teaching practice, and student achievement in middle grades mathematics. *Educ Stud Math* (76), 141-164.
- Van de Grift, W. J. C. M. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement*, 25 (3), 295-311. <https://doi.org/10.1080/09243453.2013.794845>
- Van der Werf, M. P. C., Creemers, B. P. M., De Jong, R., & Klaver, L. (2000). evaluation of school improvement through an educational effectiveness model: The case of Indonesia's PEQIP project. *Comparative Education Review*, 44 (3), 329-356.
- Van der Werf, M. P. C., Opendakker, M-C., & Kuyper, H. (2008). Testing a dynamic model of student and school effectiveness with a multivariate multilevel latent growth curve approach. *School Effectiveness and School Improvement*, 19 (4), 447-462.
- Van de Grift, W. J. C. M. (2014). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement*, 25 (3), 295-311. <https://doi.org/10.1080/09243453.2013.794845>
- Van der Lans, R. M., van de Grift, W. J. C. M., & Van Veen, K. (2017). Individual differences in teacher development: an exploration of the applicability of a stage model to assess individual teachers. *Learning & Individual Differences*, 58, 46-55. <https://doi.org/10.1016/j.lindif.2017.07.007>.
- Wim J.C.M. van de Grift, Seyeoung Chun, Ridwan Maulana, Okhwa Lee & Michelle Helms-Lorenz (2017) Measuring teaching quality and student engagement in South Korea and The Netherlands, *School Effectiveness and School Improvement*, 28 (3), 337-349, DOI: 10.1080/09243453.2016.1263215

Words: 5824

Characters: 39 843 (22,1 standard pages)

Siti Nurul Azkiyah, M.Sc., Ph.D
 UIN Syarif Hidayatullah Jakarta, Indonesia
 Jl. Ir H. Juanda No.95, Ciputat, Kec. Ciputat Tim., Kota Tangerang Selatan,
 Banten 15412,
 Indonesia
 azkiyah@uinjkt.ac.id

Prof. Amirul Mukminin, Ph.D.
 Universitas Jambi, Indonesia
 Jln. Ir. H. Juanda, Simpang 3 Sipin, RT.29/No.84,
 Kota Baru, Jambi, 36126
 Indonesia
 amirul.mukminin@unja.ac.id