

Making sense the teacher evaluation with the value of local culture (The voice of school principals in Indonesia)

Supadi Supadi

Educational Management Department, Universitas Negeri Jakarta, Jakarta, Indonesia

Evitha Soraya and Hamid Muhammad

Universitas Negeri Jakarta, Jakarta, Indonesia, and

Nurhasanah Halim

English Language Education, STKIP Kusuma Negara, Jakarta, Indonesia

Abstract

Purpose – The voice of school principals represents the principals' thoughts and experiences because of their as teachers' evaluator. It provides principals' perception on making sense the teacher evaluation. In qualitative research, voice can provide the truth and meaning of principals' experience in teachers evaluation. Their voices in the qualitative interviews are recorded and transcribed into words (Jackson and Mazzei, 2009 and Charteris and Sardon, 2018). By listening to the voices of principals in five provinces in Indonesia, this study, a qualitative research, intends to explore the principals' sensemaking in teacher evaluation.

Design/methodology/approach – This study adopted a qualitative approach, as it was principally concerned with capturing participants' direct experiences in their natural setting as both the teachers' evaluator and school leader (Patton, 2002). The qualitative interview and content analysis were used in this study. The qualitative interview is a type of conversation used to explore informants' experiences and interpretations; in this study, the headmaster (Mishler, 1986; Spradley, 1979 in Hatch, 2002). Researchers used the interviews to uncover the structure of meaning used by principals in making sense the policies that determine teacher evaluations and that are used to carry out evaluations within principal's local authority. The implicit structure can be discovered from direct observation, and the qualitative interviews can bring this meaning to the surface (Hatch, 2002). Therefore, by applying the qualitative interviews, it is expected that information or "unique" interpretations from the principal can be obtained (Stake, 2010). Content analysis is a research technique for making valid conclusions from oral texts into a research context. This analysis can provide new insights, improve researchers' understanding of certain phenomena, or inform other practical actions through the use of verbal data collected in the form of answers to open interview questions (Krippendorff, 2004).

Findings – There are three important findings relating to principals' sensemaking of teachers' evaluation; they are teachers' length of service, principals' perceptions of teacher evaluations and consistency in teacher performance improvement. The principals' perception greatly influences their beliefs and sensemaking of teacher evaluation. In essence, teacher evaluation has not been used to identify high-quality teachers. Principals focus more on the improvement of teachers' welfare than teacher actual performance.

Research limitations/implications – Future research should explore principals' attitude toward the stakeholders when student achievement is not in line with the consistent increase in teachers' performance ratings. And, it is also necessary to investigate the policy makers response to see the consistent improvement in teacher's evaluation is not in line with student achievement. Finally, how to eliminate the culture of joint responsibility without causing frictions in the school environment.



Originality/value – The authors hereby declare that this submission is their own work, and to the best of their knowledge, it contains no materials previously published or written by another person, or substantial proportions of material that have been accepted for the award of any other degree or diploma any other publishers.

Keywords Teachers' evaluation, Local culture, The voice, Principals' perception, Indonesian school

Paper type Research paper

1. Introduction

Recently, the key issue of teacher evaluation was about the belief, perceptions, sensemaking and challenges of principals as the evaluator (Kraft and Gilmour, 2016; Reid, 2017; Dodson, 2017; Shaked, 2018). Aside from the unforeseen effects that undercut their evaluation quality, there are also four principals' challenges in teacher evaluation (Kraft and Gilmour, 2016). First, differences in principals' perspectives on how they implement the evaluation; second, teacher evaluation is time-consuming; three, feedback provision needs specific expertise; and four, principals need certain training to implement. It was also perceived that principals played a more tolerant role when evaluating teachers. They paid more attention to such frequent observable aspect as teachers' attendance instead of evaluating their instruction (Darling-Hammond *et al.*, 1983; Hallinger *et al.*, 2014; Wise *et al.*, 1985, in Reid, 2017; Derrington and Campbell, 2018). This is typically because principals are concerned about their teachers' future employment (Reid, 2017).

It is ubiquitous that teacher evaluations failed to identify high-quality teachers (Papay, 2012; Reid, 2017; Shaked, 2018). Obstacles are unavoidable in carrying out effective teacher evaluation. The obstacles are (1) principal's lack of time in assessing teachers, (2) principal's lack of knowledge and skills to expand their roles of being instructional leaders, (3) principal's discomfort because the results of the evaluation could result in the dismissal of the teacher and (4) principal's preference for preventing high levels of teacher turnover (Shaked, 2018). It is presumed that time and expertise have been the challenges and obstacles in implementing the quality teacher evaluation.

Furthermore, principals would carry out teacher evaluations to bring benefits for better principal-teacher relationship and to be relevant with principals' abilities and experiences. By these, principals could prevent frequent teacher's turnover and considerable time allocation in the evaluation (Lavigne and Olson, 2019). On the contrary, the benefits and relevance lead the principals to give teachers a higher rating than what they deserve. So, even if students' achievement was relatively low, teacher ratings always reflected effective or very effective performance. Yet, unlike others, Wind *et al.* (2019) found a pattern in giving teachers' ratings determined by teacher's demographic factors.

The voice of school principals in this study represents the principals' thoughts and experiences because of their role as teachers' evaluators. It provides principals' perception of making sense teachers' evaluation. In qualitative research, the voice can provide the truth and meaning of principals' experience in teachers' evaluation. Their voices in the qualitative interviews are recorded and transcribed into words (Jackson and Mazzei, 2009 and Charteris and Smardon, 2018). By listening to principals' voices in five provinces in Indonesia, this qualitative research explores the principals' sensemaking the teachers' evaluation.

1.1 Teacher evaluation in Indonesia

In Indonesia, teacher evaluation is carried out by applying and making sense of the Guidelines in Implementing Teacher Performance Assessment issued by the Ministry of National Education. It is to create professional teachers with excellent services. A professional teacher is recognized by the possession of an academic qualification, competencies, the educator certificate, a sound body and mind, and the abilities to realize the goal of national education (Law No. 14/2005 on Teachers and Lecturers). To earn the certificate, the teacher should possess

the required academic qualification and professional teaching competencies. The certificate is obtained from a short course held by accredited teacher education programs as determined by the government.

Moreover, the guidelines stipulate the National Academic Qualification Standards and Teacher Competencies nationwide. Based on this document, teachers are professional educators with specific tasks, functions and roles in realizing the national education goals. Teachers' main tasks refer to their mastery of content knowledge and their application within and across the subject. There are four main competencies (pedagogical, personal, social and professional competencies) that are developed as teacher performance indicators. It is expected by determining the required teachers' competencies, a quality learning process can be achieved. Each competency is elaborated in varied numbers of indicators. The teacher performance assessment is designed to measure the level of competencies, knowledge and skills in task accomplishment.

Teachers are expected to engage in realizing the youths who stay devoted to God, excel in science and technology, aesthetics and are ethical with noble characters and personalities. Therefore, to measure their engagement, the teacher evaluation and teacher continuous professional development are stipulated. While teacher evaluations aim to ensure the quality of the learning process in class at all levels of education, teacher development refers to one of the government programs in developing teacher competencies. A teacher is eligible to join the professional development only after he/she has undergone the teacher's evaluation. In essence, both teacher evaluation and development are important, related and carried out in a sustainment plan.

Teacher evaluation is critical in Indonesia. It is used to monitor and improve teachers' performance to provide quality education for students. The evaluation is conducted twice a year, at the beginning of the school year (formative assessment) and at the end of the school year (summative assessment). Formative assessment is used to compile a teacher's performance profile and must be done within six weeks at the beginning of the school year. Formative assessment is in the form of a teacher's self-assessment. The results will be the basis for the teacher's summative assessment. Summative assessments are used to determine teachers' credit scores at each school and monitor the teacher's performance progress based on his/her self-assessment (formative test). The results will be the basis for teacher promotion and continuous professional development. When conducting teacher evaluations, three requirements must be fulfilled: valid, reliable and practical.

The guidelines also show three aspects that must be assessed; they are (1) the teaching process, (2) counselling and (3) additional tasks relevant to school functions. Each aspect has its components that show the standards of teacher academic qualification and competencies. The teaching and counseling process reflects the implementation of the teaching and learning process, including activities such as planning and implementing teaching and learning, assessing and evaluating students, analyzing students' performance for the follow-up activities or assessment when necessary. The final aspect is that the additional tasks relevant to the school function and time basis, i.e. during school hours or after school hours. There are four areas of competency assessed, namely, pedagogy, personality, social and professional. Each aspect also has indicators to be appraised.

Concerning the challenges and obstacles discussed earlier, the school principals in Indonesia might encounter similar issues in teacher evaluation. The considerable number of indicators to be appraised might reflect that time allocation must have been the first challenge. Principals must dedicate considerable time to implement the evaluation. It suggests that school principals might prefer to give higher ratings than to rate based on what teachers deserve to save time. As a matter of fact, the school principal has many other responsibilities to fulfill as the school manager. Also, special skills and experiences are needed to provide positive and effective feedback.

2. Research design

Qualitative methods offer an effective way of capturing participants' direct experiences in their natural setting as both the teachers' evaluator and the school leader (Patton, 2002). The qualitative interview and content analysis were used in this study. The qualitative interview is a type of conversation used to explore informants' experiences and interpretations; in this study, the principal (Mishler, 1986; Spradley, 1979 in Hatch, 2002). Researchers used the interviews to uncover the structure of meaning used by principals in making sense of the policies that determine teacher evaluations and that are used to carry out evaluations within the principal's local authority. The implicit structure can be discovered from direct observation, and the qualitative interviews can bring this meaning to the surface (Hatch, 2002). Therefore, by applying the qualitative interviews, it is expected that information or "unique" interpretations from the principal can be obtained (Stake, 2010).

Content analysis is a research technique for making valid conclusions from oral texts into a research context. This analysis can provide new insights, improve researchers' understanding of certain phenomena or inform other practical actions through verbal data collected in the form of answers to open interview questions (Krippendorff, 2004).

2.1 Participants

Of 150 elementary and secondary school principals contacted, 50 principals (37 females; 13 males) agreed to participate in this study. They were 50–60 years old, and most of them (60%) held a master's degree in education. Afterward, they were asked to complete a consent form and briefed about the research purposes and benefits. These principals were purposively recruited for the following criteria: (1) they had served as school principals for more than ten years; (2) they had sufficient training in educational management and teacher evaluation; and (3) they had experiences in managing fully- and under-resourced government-owned schools under the supervision of the Ministry of Education and Culture. The purposive random sampling technique was used to maintain the credibility of the results (Patton, 2002) and to be able to capture the perceptions of principals from various angles to obtain detailed information (Merriam, 2009, in Cresswell, 2012). Sampling is deliberately chosen based on the standards used in selecting participants and sites whether they are "rich in information" (Patton, 1990 in Cresswell, 2012). Maximum variation sampling is carried out in this study by considering the principal's experience in evaluating teachers, school level (elementary and secondary) and school location (urban). In this study, the selected school locations were Jakarta, West Java, Banten, Yogyakarta and Solo. The selection could represent the sensemaking of principals in the urban areas of Indonesia.

2.2 Data collection procedure

The qualitative interviewing was conducted by asking school principals who became the samples individually (one-on-one interview to explore in-depth information about how school principals evaluated teachers). When the consent forms had been completed, the school principals agreed to have multiple interviews. Interview schedules were negotiated and done at the school principal's convenience. The interview began with the assumption that the principals' perception is meaningful, collected and interpreted explicitly. Interview guides were also prepared and used to ensure a similar basis when questioning samples during the interview. The guides made the interview more systematic and comprehensive by limiting the problem to be explored (Patton, 2002). The questions asked were open-ended, making the interviewer free to build a conversation to utilize the available time. Interviews were also conducted in an atmosphere that did not make the principal feel intimidated. With permission, all the interviews were audio-recorded during the interview. Then, the interviews were transcribed so that the meaning-making can be done. The core consistency and meaning can be identified (Patton, 2002; Cresswell, 2012). Using one-on-one interviews was quite

time-consuming and costly, but there is no doubt of the data completeness. Meanwhile, content analysis is applied to reveal the principals' perception (Boyatzis, 1998, in Patton, 2002).

2.3 Data analysis

Data analysis was carried out through four stages: condensation, coding, categorization and theorizing. After the data were collected, condensation was performed by sorting principals' utterances relevant to the principal's consideration in giving ratings to the teacher (Miles *et al.*, 2014). Second, the transcribed interviews were coded based on the principals' system, aspects or thoughts. Codes are words or phrases that symbolically provide summative, prominent, capturing essence and/or evocative attributes for some language-based or visual data (Saldana, 2015). In qualitative data analysis, codes are constructs developed by researchers, and coding is considered a "critical link" between data collection and explanation of its meaning (Charmaz, 2001 in Saldana, 2015). After capturing the essence of utterances, similar utterances are grouped to generalize meaning and obtain categories in the third stage. Finally, the theorizing stage was performed. This stage aims to achieve the conceptual constructs of categories obtained in the previous stages and see how they are interconnected and how they influence each other as part of an abstract construct (Richards and Morse, 2013, in Saldana, 2015).

In this study, the representativeness of sampling, data collection and data analysis was validated by two experts of educational management and qualitative research. Validation was also done by asking participants to examine and determine whether researchers have accurately reported their perceptions or also called *member validation method*.

3. Findings

At first, there are three codes/categories constructed in this study, i.e. (1) demographic factors that influence the evaluation, (2) principals' perceptions of teacher evaluations and (3) consistency in teacher performance improvement. Upon the theorizing stage, another code was constructed to connect each category.

3.1 Teacher demographic factors in teacher evaluation

During interviews, some principals (55%) mentioned one demographic factor influencing teacher evaluations, i.e. differences in teacher's length of service, or seniority. An interviewee from Cirebon said: "The given rating illustrates seniority." An interviewee from Banten put it as: "When rating teacher in evaluation, seniority is the first basis . . ." It was also mentioned by another interviewee from Serang that: ". . . senior teachers must be rated higher than the junior. . ."

However, the majority of interviewees (63%) did not perceive seniority based on the length of the teacher's service only. Seniority refers to the job grade in civil service and the possession of educator's certificate. The utterances were as follows: "the judgment we made is based on job grade"; "When the job grade of all teachers have been listed, we can get the highest and lowest rank"; "In determining the rating, senior teacher and the possession of educator's certificate were the principal's first considerations"; and "The way we determine the rank is by making a descending rank of teachers. Those who have served for the longest period will be in the highest ranks. We continued listing them until we got the teacher with the shortest period of service."

The finding implies that there is a classification of the teachers, which is based on their status of employment, level of job grade, possession of educator's certificate and length of service. This classification serves as the foundation to determine the scale of rating a teacher could deserve and to decide whom of the senior teachers gets the highest rating. The following are the statements:

- (1) "I think it is fair to rate senior teachers higher as they have served for such a long period of time."
- (2) "We listed the teachers based on their employment status and the length of service, from the longest to the shortest period of service. The list will start from full-time teachers with permanent employment status and the longest period of service to the shortest. Then, we listed the part-time teachers from the longest to the shortest period of service."
- (3) "The rank used to determine teacher's rating is actually based on seniority. The higher the rank, the longer period of service. We took into account considerable teaching experience of junior teachers and the full-time teachers with permanent employment have become the main consideration in determining the rating."

In short, seniority is characterized by higher job grade, employment status, possession of educator's certificate and length of service. Seniority significantly influences a principal's perception in teacher evaluation. In fact, this is the only demographic factor that the principal considers when evaluating a teacher. The finding presented in this section suggests that senior teachers receive a higher rating than they deserve.

3.2 Principals' perception of teacher evaluations

In this study, participants perceived teacher evaluations with three aspects: the fulfillment of teaching resources (48%), evaluation of teacher's competency (40%) and teacher's academic qualification standards (12%).

Figure 1 presents the principals' perception, which partly reflected the essence of teacher performance evaluation, or teacher evaluation based on the National Education Ministry Guidelines. There was a common view among interviewees about this. The following is the summary: "There are four domains of teacher's competencies: pedagogical competency, social competency, personality competency, and professional competency. The principal assesses pedagogical competency based on how the teacher knows and understands his students, the teacher designs and develops the teaching programs, and they develop students' competencies. The principal assesses the teacher's social competency by assessing how an individual teacher interacts with their colleagues and principals. Next, personality competency is assessed from whether the teacher has been a good role model for his/ her students. Finally, the teacher's professional competency is assessed based on the teacher's classroom management and the teacher's subject mastery and handling of class questions." Based on this, the fulfillment of teaching resources and evaluation of teacher's competencies

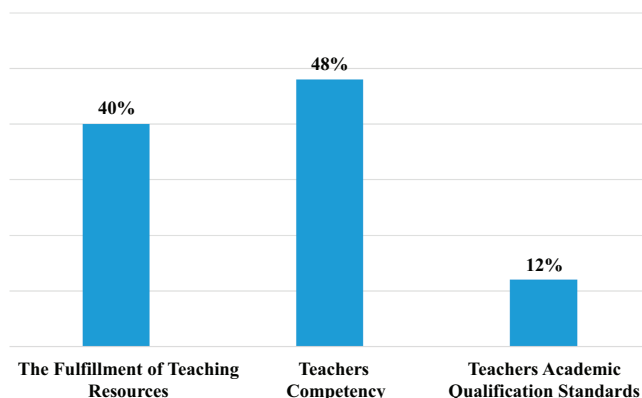


Figure 1. Principals' perceptions of teacher evaluations based on frequency of utterances

reflects the teacher's pedagogical and teacher's professional competencies. While, teacher's academic qualification standards are indicated from the teacher's social and personality competencies. Given all that has been mentioned so far, principals should have used four competencies to identify high-quality teachers through teacher evaluation.

On the other hand, from the principals' voice, there are several aspects included in the principals' sensemaking when evaluating teachers. They are discipline, level of attendance, dedication and professionalism of teachers. Discipline was the representation of teacher's accuracy in starting and ending the teaching process. The teacher should focus entirely on teaching students. The level of attendance was taken into account to prevent teacher's counterproductive such as skipping classes without prior notice to avoid fulfilling his obligations, giving assignments without in-class explanation, etc. A principal assesses a teacher's dedication by measuring the teacher's ability to complete his/her tasks on time, use his/her free time for productive activities and measure teacher's contribution to school progress and achievement. Referring to the Guidelines of Implementing Teacher Performance Assessment, there are three indicators to measure this competency: (1) a teacher act according to Indonesian law, (2) teachers present themselves as role models for students and society, (3) and teachers behave in accordance with the professional code of ethics for teachers. To confirm the findings, the four aspects are especially relevant with the last indicator only.

Some participants also perceived responsibility, loyalty, honesty and empathy are indicators relevant to teacher evaluation. In the guidelines, this is relevant to personality competency. It derives from the explanation of the first indicator "teacher respects and promotes the Five principles of Indonesians (Pancasila) as the ideological and ethical bases for all Indonesian citizens, and teacher respects each other." Moreover, some participants perceived teachers' professionalism as teachers' readiness to prepare their learning media, attendance sheet, lesson plan and teacher's notes during class. However, referring back to the previous paragraph, teacher's preparation for teaching resources indicates the teacher's pedagogical competency. In this competency, the teacher is assessed through his abilities to develop effective lesson plans which fulfill students' needs based on their characteristics, to implement the plan in teaching/learning activities, to compile and using various learning materials and resources, and if possible, to take the benefits of using information and communication technology (ICT) within the teaching learning process. On the other hand, the attendance sheet and teacher's notes in class are part of a teacher's professional competency.

Based on the guidelines, teacher effectiveness is assessed by comprehensive indicators such as the teacher can track and frame students' performance through the identifications of students' task accomplishment and difficulties, the lesson plan development, the efficiency of implementing the lesson plan based on the time allocation, materials selection and preparation. Taken together, the results suggest that principals' perceptions of teacher evaluations are in line with the Guidelines for Implementing Teacher Performance Evaluation. However, principals tend to focus more on teacher competencies that are visible and observable than measurable competencies. Of the four competencies evaluated, personality and professional competencies were more dominant in shaping the principal's perception of teacher evaluation. From these findings, the researchers suggest the need for more comprehensive training, seminars or discourse for school principals so that principals' sensemaking in teacher evaluation could cover teachers' pedagogical and social competencies.

3.3 Consistent improvement in teacher performance

Teacher evaluation is intended to obtain an individual teacher's performance profile. Almost all participants (85%) implied that the teacher evaluation had not compiled the teacher performance profile. In the interview results, two participants clearly stated this. These are the statements:

- (1) To be honest, we implemented the teacher evaluation only to fulfill the principal's duties.
- (2) We rated teachers by considering their job grades; therefore, we have not paid attention to their actual performance.

If the results of teacher evaluations have failed to represent teacher's actual performance, it can be concluded that the teacher evaluation might not be sufficient in portraying teacher performance profile. In other words, the principal has rated the teacher higher than what he deserves. Furthermore, it happens to all teachers. This leads to the question of how to rate the teacher only based on their length of service.

Taking seniority as a consideration in teacher evaluation also shows teachers' acceptance of junior teachers will be rated lower than the senior ones. As teacher evaluation is used for teacher's promotion, school principals perceived that taking the teacher's length of service as the basis to rate the teacher is fair. This is in line with [Wind et al. \(2019\)](#), who revealed there is a pattern in rating teachers. In this study, it was found that the pattern was to put teacher length of service in the descending list. The list will include the teacher's status of employment, length of service and job grade. This is the answer to how to rate teachers based on seniority. Here are some relevant statements taken from the interviews:

- (1) "The way I took in determining the rating was by holding an informal meeting with all teachers, a consensus. We agreed that the junior teachers would assist senior teachers in preparing the lesson plan, learning media, and other requirements, and junior teachers accept that senior teachers are rated higher than they deserve."
- (2) "I invited all teachers in an informal meeting to agree that all teachers must get the promotion."
- (3) "It is no longer a secret to assist teachers in getting promotion and better earnings. (Therefore), teachers are rated higher than what they deserve."

3.4 Regulation or harmony of principal–teacher relations?

With the demand to become instructional leaders, principals are expected to improve their teacher's performance (Glick, 2011 in [Derrington and Campbell, 2018](#)). This implies that there must be harmony in the relationship between the principal and the teachers. In the interview, participants also stated that the principal gave senior teachers a higher rating to maintain the harmonious relationship between the principal and teacher. The research participants expressed it as follows:

- (1) "... the rating was higher, so it does not reflect the actual performance. As long as the teachers are senior, even their performance was low, we must rate them higher than what they deserve to maintain a harmonious relationship between school members."
- (2) "... and to avoid frictions among teachers and to maintain a good relationship with the principal, the rating is somehow higher or similar to the previous evaluation."
- (3) "... the point is the rating should be higher than the previous evaluation. This is to create a positive school climate among teachers and school principals ..."
- (4) "... so the principal never evaluates the teacher's actual performance, to avoid gaps between senior and junior teachers."

It was also found in the interview that government regulation has caused this to happen. The regulation requires a consistent increase in teacher performance evaluation. This was stated by participants as follows:

- (1) “The rating must be higher than the previous evaluation; therefore, the principal increases the rating.”
- (2) “In essence, when I have to evaluate teachers, I need to look at the last year’s evaluation record. Then, I increase the rating a little, so yes, the teacher evaluation does not reflect the teacher’s actual performance.”
- (3) “. . . low performance is not acceptable, so the rating should be high.”

Later from interviews in Jakarta, it was found that there is no such regulation as principal must consistently increase teacher’s rating but rather what is commonly called “joint responsibility” in a business context. The following statement was collected from a participant who was a principal in junior high school:

- (1) “. . . It is a joint responsibility, meaning that when our subordinates made errors or had obstacles, perhaps their performance was low, the superior officers will also bear the consequences. So, it is true that the performance of teachers has potentially affected the principal.”

It should be concluded that when the principal gives a teacher higher rating than what he/she deserves, it is a shared responsibility of what the teacher has performed in teaching students. It is considered the principal’s anticipative action toward greater problems or bad impact due to the low rating in teacher performance. To sum, it reflects the principals’ tolerant attitude toward teacher evaluations.

4. Discussion

Returning to the question posed at the beginning of this paper, it is now possible to state that a principal’s perception greatly influences his beliefs and sensemaking of teacher evaluation in Indonesia. The determinant of principals’ perception is one teacher demographic factor, i.e. length of teaching service or seniority. In the sensemaking of teacher evaluations, (1) ranking teachers using a pattern of ascending list based on seniority, (2) rating teacher based on measurable aspects instead of observable aspects, (3) rating senior teachers higher than what they deserve and make this as an agreement with all teachers and (4) prioritizing the harmony of principal–teacher relationship and teacher–teacher relationship are employed. Unfortunately, the principal’s sensemaking of implementing the teacher evaluation covers only two of four competencies assessed in the teacher performance evaluation, i.e. personality competency and professional competency. It is highly recommended that principals’ sensemaking were extended through training or short courses.

The biggest challenge of school principals in Indonesia is the principal’s perspective on how principals implement the evaluation (Kraft and Gilmour, 2016). The principal believes that it is acceptable to be tolerant in evaluating teachers. This is because the principal’s concern for teachers under his supervision to get promotion and pursue their professional development in the future. The evaluation is implemented by considering the consistent improvement of teacher performance and the harmony in principal–teacher relations. Hence, reporting consistent improvement of teachers’ performance without considering their actual performance is not a wise choice, but maintaining a positive school climate is far more important (Marraccini *et al.*, 2020). This situation has confirmed what Flores and Derrington (2017) call conflicting goals.

Although there have been considerable indicators in each domain of assessed competencies, it is interesting to note that the principal did not perceive the length of time to implement the evaluation has not been a challenge or an obstacle. As the local autonomy owner, the principal uses his authority to evaluate teachers by adapting the recommended procedures. This shows that principals never eliminate teacher performance evaluations, but

they have been more tolerant in rating them during evaluation. School principals' perceived teachers' planning for effective teaching and learning process, and teachers' fulfilling the academic qualification standards should focus on attention to indicate teachers' quality. School principals also perceive that evaluating observable competencies is more beneficial than the measurable ones. Hence, the principal shall provide direct feedbacks and have the opportunities to interact with teachers. Referring to the research findings, it can be inferred that principals focused more on the formative assessment. It means teacher evaluation is used only to compile a teacher's performance profile. Later, the assessment results will be used by school principals as a basis to propose the teacher to join a certification program or future professional development programs such as training, seminars, workshops, etc. In short, teacher performance evaluation has not been optimally employed.

Another thing to note is that teacher–student relationship was not relevant to the principal's perception of teacher's evaluation. The teacher–student relationship does not affect the harmonious relationship between the principal and the teacher.

Based on the research finding, the principal has been tolerant in teacher evaluation by assessing discipline, level of attendance, dedication, responsibility, loyalty, honesty and empathy, instead of the guidelines' prescribed aspects. Principals have expanded their autonomy in understanding and implementing teacher evaluation in a specific way (Reid, 2017). It does not happen in Indonesia only, but this has become a phenomenon in the USA as teacher evaluation has been the basis for teachers to be promoted (Papay, 2012). Furthermore, principals have also expanded their role from school managers to instructional leaders who support teacher development through the evaluation process (Kraft and Gilmour, 2016).

The need for training to provide useful feedback was not in the principal's interest to implement a valid, reliable and practical teacher evaluation. Feedback is given in the context of verbal communication based on direct observation. It seems that principals never provide written feedback because the principal always tries to solve emerged problems. With this persuasive approach, the principal expects teachers to push themselves to be more competent better so that ratings reflect their actual performance.

An additional question that has arisen when analyzing the findings is whether the implemented teacher evaluation system has fulfilled its three requirements: valid, reliable and practical. According to researchers, the three requirements refer to the instruments used in the evaluation itself. Furthermore, because the instruments used are under those stipulated in the guidelines, the implementation of teacher evaluation has fulfilled the requirements.

In Indonesia, the first challenge in teacher evaluation ineffectiveness is that principals always consider teacher evaluations should bring benefits for principals' and teachers' harmonious relationship. If it is considered to cause a gap between the teacher and principal, the principal tries to embrace all teachers through agreement.

Yet, the biggest challenge in evaluating teachers in Indonesia is the culture of joint responsibility. This culture is widely used in economics to indicate a shared responsibility in bearing the risk of a form of joint investment. The culture may be developed from the value of mutual assistance (*gotong royong*) – the key element in the Indonesian system of political and cultural power. The spirit is to bring together everyone in a community to be in a mutual tolerance to take responsibility for improving teachers' quality and welfare. This culture has led principals' inconvenience to give low ratings even when teachers deserve it. The principals' concern is to improve teacher welfare through promotion by rating teachers higher than what they deserve.

5. Conclusion

It can be concluded that the results of teacher evaluation in Indonesia have not been able to identify teacher performance. Therefore, the teacher evaluation cannot be used as the

standard to assure the quality learning process in the classroom. From the theories' standing, teacher evaluation has not been used to identify high-quality teachers in Indonesia and to analyze teacher's actual performance. Teacher evaluation is effective because teachers understand and develop their self-assessment indicators at the beginning of each semester (formative assessment). The principal's authority is to confirm teacher performance progress based on the self-assessment by rating the indicators at the end of the academic year (summative assessment). Unfortunately, teacher evaluation only focuses on formative assessment, so the teacher's progress in their teaching performance has not been measured. In other words, teacher evaluations' employment is only for reference to support teacher promotion and teacher professional development. Principals, as the local autonomy, focus more on the improvement of teacher welfare through teacher evaluation.

From the practical implications, the current research findings indicate that to establish more effective teacher evaluation policies, perceptions, challenges and obstacles when carrying out teacher evaluations should be considered. It was found that ratings given by school principals varied based on seniority, which refers to the length of teaching service, employment status, level of job grade and possession of educator's certificate. To determine the high and low teacher's rating will be based on these three aspects. For example, teachers who have a longer teaching service period, full-time teachers and certified teachers will get higher ratings than teachers who only have one or three of these things. The rating does not reflect the actual performance. Thus, teacher evaluation results are not in line with the classroom's quality of learning or student achievement.

Future research should explore the principal's attitude toward the stakeholders when students' achievement is not in line with the consistent increase in teachers' ratings. Furthermore, it is also necessary to investigate the policymaker's response to the consistent improvement in teacher evaluation, which is not in line with students' achievement. Finally, it is necessary to eliminate the culture of joint responsibility, without causing frictions in the school environment.

References

- Charteris, J. and Smardon, D. (2018), "Assessment and student participation: 'choice and voice'", in *School Principal Accounts of Schooling Territories, Teaching Education*. doi: [10.1080/10476210.2018.1462311](https://doi.org/10.1080/10476210.2018.1462311).
- Creswell, J.W. (2012), *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*, 4th ed., Pearson, Boston.
- Derrington, M.L. and Campbell, J.W. (2018), "High-stakes teacher evaluation policy: US principals' perspectives and variations in practice", *Teachers and Teaching*, Vol. 24 No. 3, pp. 246-262, doi: [10.1080/13540602.2017.1421164](https://doi.org/10.1080/13540602.2017.1421164).
- Dodson, R. (2017), "An analysis of principal perceptions of the primary teaching evaluation system used in eight U.S. States", *International Journal of Education Policy and Leadership*, Vol. 12 No. 5.
- Flores, M.A. and Derrington, M.L. (2017), "School principals' views of teacher evaluation policy: lessons learned from two empirical studies", *International Journal of Leadership in Education*, Vol. 20 No. 4, pp. 416-431, Retrieved March 6, 2020, doi: [10.1080/13603124.2015.1094144](https://doi.org/10.1080/13603124.2015.1094144).
- Hatch, J.A. (2002), *Doing Qualitative Research in Education Settings*, State University of New York Press, New York.
- Jackson, A. and Mazzei, L. (Eds) (2009), *Voice in Qualitative Inquiry: Challenging Conventional, Interpretive, and Critical Conceptions in Qualitative Research*, Routledge, London, pp. 97-113.
- Kraft, M.A. and Gilmour, A.F. (2016), "Can principals promote teacher development as evaluators? A case study of principals' views and experiences", *Educational Administration Quarterly*, Vol. 52 No. 5, pp. 711-753, Retrieved March 6, 2020, doi: [10.1177/0013161X16653445](https://doi.org/10.1177/0013161X16653445).

-
- Krippendorff, K. (2004), *Content Analysis: An Introduction to its Methodology*, 2nd ed., Sage Publications, California.
- Lavigne, A.L. and Olson, A.M. (2019), "Teacher evaluation through the eyes of the principal: how individual and school contexts shape perceptions of practice", *Leadership and Policy in Schools*, Vol. 19 No. 4, pp. 575-596, doi: [10.1080/15700763.2019.1637899](https://doi.org/10.1080/15700763.2019.1637899).
- Marraccini, M., Fang, Y., Levine, S.P., Chin, A. and Pittleman, C. (2020), "Measuring student perceptions of school climate: a systematic review and ecological content analysis", in *School Mental Health*, Vol. 12, pp. 195-221, doi: [10.1007/s12310-019-09348-8](https://doi.org/10.1007/s12310-019-09348-8).
- Merriam, S.B. (2009), *Qualitative Research: A Guide to Design and Implementation*, John Wiley & Sons, San Francisco.
- Miles, M.B., Michael Huberman, A. and Saldaña, J. (2014), *Qualitative Data Analysis: A Methods Sourcebook*, 3rd ed., SAGE Publications, Los Angeles.
- Papay, J.P. (2012), "Refocusing the debate: assessing the purposes and tools of teacher evaluation", *Harvard Educational Review*, Vol. 82, pp. 123-141, doi: [10.17763/haer.82.1.v40p0833345w6384](https://doi.org/10.17763/haer.82.1.v40p0833345w6384).
- Patton, M.Q. (2002), *Qualitative Research and Evaluation Method*, SAGE, California.
- Reid, D.B. (2017), "U.S. Principals' interpretation and implementation of teacher evaluation policies", *The Qualitative Report*, Vol. 22 No. 5, pp. 1457-1470, Retrieved March 6, 2020, available at: <https://nsuworks.nova.edu/tqr/vol22/iss5/19>.
- Saldana, J. (2015), *The Coding Manual for Qualitative Researchers*, 3rd ed., SAGE Publications, Arizona.
- Shaked, H. (2018), "Why principals often give overly high ratings on teacher evaluations", *Studies in Educational Evaluation*, Vol. 59, pp. 150-157, Retrieved March 10, 2020, doi: [10.1016/j.stueduc.2018.07.007](https://doi.org/10.1016/j.stueduc.2018.07.007).
- Stake, R.E. (2010), *Qualitative Research: Studying How Things Work*, The Guilford Press, New York.
- Wind, S.A., Jones, E., Bergin, C. and Jensen, K. (2019), "Exploring patterns of principal judgments in teacher evaluation related to reported gender and years of experience", *Studies in Educational Evaluation*, Vol. 61, pp. 150-158, Retrieved February 20, 2020, doi: [10.1016/j.stueduc.2019.03.011](https://doi.org/10.1016/j.stueduc.2019.03.011).

Further reading

- Bergin, C., Wind, S.A., Grajeda, S. and Tsai, C.-L. (2017), "Teacher evaluation: are principals' classroom observations Accurate at the Conclusion of training?", *Studies in Educational Evaluation*, Vol. 55, pp. 19-26, Retrieved May 12, 2020, doi: [10.1016/j.stueduc.2017.05.002](https://doi.org/10.1016/j.stueduc.2017.05.002).
- Bowen, J.R. (1986), "On the political construction of tradition: Gotong Royong in Indonesia", *The Journal of Asian Studies*, Vol. 45 No. 3 (May), pp. 545-56, Retrieved June 6, 2020, doi: [10.2307/2056530](https://doi.org/10.2307/2056530).
- Koelsch, L.E. (2013), "Reconceptualizing the member check interview", *International Journal of Qualitative Methods*, Vol. 12 No. 1, pp. 168-179, Retrieved March 6, 2020, doi: [10.1177/160940691301200105](https://doi.org/10.1177/160940691301200105).

Corresponding author

Nurhasanah Halim can be contacted at: nurhasanah_halim@stkipkusumanegara.ac.id

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com