



Teacher's Performance Appraisal Practice and Challenges in Government Primary Schools of Dilla City Administration

Befkadu Legesse Abate *

* Department of Educational Planning and Management (PhD Candidate), Dilla University, Ethiopia

Received: 12 January 2023

Accepted: 12 March 2023

Published: 09 April 2023

ARTICLE INFO.

Key words/phrases:

Performance appraisal,
Performance appraisal practice
and challenges, Teacher

Abstract

The main purpose of this study was to assess the current practice and challenges of teachers' performance appraisal practices in government primary schools under the Dilla City administration. The research utilized a mixed-method concurrent triangulation design approach. There were four complete primary schools in the city, and two of them were included in the study. From these schools, 62 teachers, two principals, one school supervisor, two department heads, two PTAs, and two students were selected as a sample. Concerning the sample, teachers were selected using proportional and systematic sampling techniques. Contrary to this, the purposive sampling technique was used to select principals, department heads, and student council; PTAs and one supervisor were also selected by the available sampling method. Findings from both instruments were presented in an integrative way. The findings of the study revealed that teacher performance appraisal has been a common practice, often conducted twice a year, but with the principal as the dominant appraiser. Concerning purpose, the current system of teachers' performance appraisal has not mainly served the developmental drives of PA. The criteria used to appraise teachers were viewed by teachers as fine, but the participation level in the formulation process was forced by the South American Nationalities Education Bureau. Concerning common challenges in the TPA: poor feedback system, low teacher participation, post-appraisal discussions, students' bias, and sometimes using PA for revenge. The text advocates for teacher involvement in formulating TPA criteria, involving stakeholders, and adapting guidelines, while also recommending meticulous assessment of teacher performance using interpersonal communication.

1 Introduction

1.1 Background of the Study

To achieve the objectives for which it was founded, an organization needs both human and non-human resources. The most valued and in-demand of these resources for the effective achievement of the intended goals is people. This is frequently the case since human resources offer the ability, skills, and efforts necessary to efficiently and effectively employ all available resources (Zeng & Qi, 2021). Any organization's management and oversight of its overall activities, whether directly or indirectly,

depends on its human resources.

Performance appraisal constitutes an integral component of human resource management that involves systematically assessing and analyzing employees' work performance, followed by constructive dialogue between the evaluator and the employee regarding the results (Ozkeser, 2019). In this sense, performance appraisal can be understood as a formalized and organized process designed to examine job-related behaviors and outcomes, identify the underlying factors influencing performance, and provide guidance for improved future effec-

tiveness in ways that create value for employees, organizations, and society at large.

It is considered an instrument through which an employee's performance and capabilities can be matched to the job requirements and career plan over time. A properly designed performance appraisal system serves as a device for better communication and development of the employee as well as for the attainment of organizational goals and objectives (Negash *et al.*, 2021). General personnel choices are made using a performance appraisal. It contributes to choices about promotions, transfers, and terminations (Islami *et al.*, 2018). It acts as a tool for determining training and development requirements as well as a standard by which programs for selection and development are evaluated.

The objective of giving employees feedback on how the organization perceives their performance is likewise achieved via appraisal. Similar to other social institutions, schools were created to support the teaching and learning process. They are institutionalized to alter and enhance kids' conduct (Monari & Wanjau, 2022). A system of teacher performance appraisal (TPA) properly designed and implemented is believed to have favorable results in the professional development of teachers. In school settings, according to Vansteenkiste *et al.* (2020), teachers are the primary expensive resources who need and deserve support and encouragement to extend their skills and the frontiers of their knowledge.

In the Ethiopian context, the current teacher evaluation framework follows a results-oriented performance appraisal (ROPA) model, which classifies performance into four levels: poor, acceptable, very good, and excellent. According to Mohammed (2020), this system assesses teachers primarily on their demonstrated achievement of key, major, and minor responsibilities assigned to them. Furthermore, within Ethiopia's primary education sector, advancement along the professional career structure is determined by both appraisal outcomes and years of service. Based on these criteria, teachers may progress through nine hierarchical ranks: beginner teacher, junior teacher, teacher, senior teacher, associate teacher, lead teacher, senior lead teacher I, senior lead teacher II, and senior lead teacher III

(SREB, 2017).

Teacher performance appraisal in Ethiopia has lots of problems. According to Shugate (2020), some of the problems with teaching performance evaluations in secondary schools in Ethiopia include inadequate evaluator skills, performance criteria that are unrelated to instructors' professions, and a negative attitude towards the entire system of performance ratings.

Presently, government schools at all levels execute performance appraisals, which are occasionally done. Nevertheless, in many of these schools, performance appraisal appears to receive formal recognition and is conducted on a routine basis, yet it functions largely as a procedural requirement rather than as a meaningful motivational mechanism through which important administrative and professional development decisions are informed.

The primary objective of this research was to examine the existing practices and associated challenges of teacher performance appraisal in government primary schools within the Dilla City Administration. In addition, the study seeks to address identified gaps in current implementation and to propose practical recommendations that may enhance the effectiveness and overall quality of performance appraisal practices in the selected schools. Therefore, at the end of this study, the researcher planned to achieve the following specific objectives:

- Examine the existing performance appraisals in those government primary schools.
- Assess the awareness of teachers about the purpose of the current TPA.
- Identify the challenges of teachers' performance appraisals in schools.

2 Research Method Approach

For this study, a mixed method was employed. A mixed-methods procedure that combines data from both quantitative and qualitative methods to provide a better understanding of a research problem. It can be useful when unexpected results arise, deepen understanding of quantitative findings, and help design and validate research instruments. Additionally, interviews and document analysis fall

under the qualitative approach (Creswell & Zhang, 2009). Mixed research methods are convenient for the social and human sciences as distinct research is used to collect extensive data and confirm findings from different data sources. In general, mixed methods research represents research that involves collecting, analyzing, and interpreting quantitative and qualitative data in a single study or in a series of studies that investigate the same underlying phenomenon.

Researchers may use various mixed-methods research designs depending on the study topic, available data, and available resources (Dawadi *et al.*, 2021). For this study, the “researcher used convergent parallel triangulation mixed methods designs. This design benefits the researcher, who converges or merges quantitative and qualitative data to provide a comprehensive analysis of the research problem” (Creswell & Zhang, 2009). In this procedure, the qualitative and quantitative strands are usually weighted equally.

Hence, the researcher used triangulation to validate quantitative data for this study. In this context, the researcher collects both types of data at the same time. In short, interview qualitative items were adjusted and held at a time during the collection period of the quantitative survey.

Under this design, findings derived from quantitative analysis were systematically examined alongside qualitative evidence to ensure confirmation and strengthen the credibility of the results. The qualitative strand enables researchers to gain deeper insight into complex realities, particularly by capturing participants’ perspectives and lived experiences. In contrast, the quantitative strand provides the analytical basis for examining relationships among measured variables, thereby supporting explanation, prediction, and, where applicable, control of the phenomena under investigation.

2.1 Population

The target population of this research was all the permanent teachers, principals, cluster supervisors, department heads, PTAs, and student councilors in the entire four primary target schools in the Dilla City administration. In Dilla city administration,

there were four full-cycle primary schools.

2.2 Sample Size and Sampling technique

Mulualem *et al.* (2022) define sampling as the process of selecting a subset of the population to serve the total number of primary schools in the city. In the two full-cycle primary schools, there were 73 teachers and 18 administrative bodies that could participate in the appraisal of the performance of teachers (six principals, four department head teachers, four PTA representatives, four representatives’ student council, and one cluster supervisor).

According to Dhivyadeepa (2015), in order to perform a stratified random sample, the researcher has to obtain a comprehensive list of the population and clearly split each unit into one stratum. Thus, the researcher took into consideration the number of teachers, department heads, and student council members while choosing a sample from a school for this study using stratified and simple random selection approaches.

With regard to the teacher sample, the required number of participants was initially determined using Cochran’s sample size determination procedure. Following this calculation, teachers were selected through proportional allocation combined with systematic sampling methods. From each of the selected schools, participants were identified systematically to ensure fairness in selection. Approximately 50% of the teachers from each school were included, resulting in a total sample of 62 teachers. Proportional representation was applied to reflect the relative size of each school’s teaching staff within the overall population. After establishing the final sample size, a systematic selection process was implemented whereby every third teacher listed on the official staff roster was chosen according to the predetermined sampling interval formula,

$$K = \frac{N}{n},$$

Where, "N" is the total number of teachers in the school and "n" is the sample required from that population.

Utilizing a formal tone, the selection of 62 teachers from the sample schools was accomplished through

the application of the formula $K = \frac{N}{n}$, where "n" signifies the necessary sample from the population and "N" denotes the entire teacher population in the respective schools. Moreover, two Parent-Teacher Associations (PTAs) were selected via purposive sampling, which entailed the active participation of two PTA members in the teacher appraisal process, while the remaining individuals did not actively participate in PTA tasks. In essence, two department heads with a year or more of experience and two student representatives from grade eight were chosen through purposive sampling based on their maturity level and academic performance. Additionally, two principals, one from each school, were selected to provide accurate data that would best enable them to address the research questions. As noted by Manna and Mete (2021), purposive sampling allows researchers to choose a sample based on their informed judgment and understanding of the population. Furthermore, one supervisor was selected using the available sampling method.

2.3 Data Collection Instruments

Considering the aforementioned fact, three types of data collection instruments were used: primary data sources, namely the questionnaire and the interview. Secondary data sources were relevant documents that the researcher used as additional evidence for the study. Questionnaires were used to collect information from teachers, and interviews were used to collect information from principals and supervisors, student council members, department heads, and PTA members. The questionnaires were composed of closed-ended items and contained demographic variables about the respondents.

The closed-ended items were treated using a 5-point Likert scale of strongly agree (SA), agree (A), moderate (M), disagree (DA), and strongly disagree (SD) with their respective values of 5, 4, 3, 2, and 1, respectively. The questionnaires contain five different parts and 34 closed-ended questions. When responding due to language barriers, the questionnaire distributed to teachers was translated from English to Amharic. Most of the items were adopted from previous related studies (Baye, 2021). An equal number of the questions were restructured by the researcher himself.

Interviews are a valuable tool for researchers, as they facilitate the collection of diverse perspectives from participants and enable the gathering of detailed, first-hand information (Kelly *et al.*, 2010). The study utilized interviews to gather valuable information from selected informants, including department heads, student council members, PTAs, school principals, and cluster supervisors, based on their roles. This study utilized semi-structured interviews to gain in-depth insights into the strategies and measures employed by each school to achieve their goals, particularly regarding teacher performance appraisal practice, due to their flexibility in question order.

The study analyzed sample schools' performance appraisal records for the 2021–2022 academic year, portfolio documents, and appraisal criteria, as well as portfolio documents demonstrating the link between appraisal and the TPA process, criteria, problems, and interventions.

2.4 Procedures for Data Collection, Analysis, and Interpretation

The initial step in data analysis was systematic data collection, which entails gathering data sets that are examined for patterns and trends (Englander, 2012). This process involves utilizing a variety of data collection techniques, including surveys, focus groups, interviews, observations, experiments, and secondary data analysis. These techniques are employed to gather data from a range of sources, investigate it for trends and patterns, and interpret it to gain new insights. In this study, data was collected through surveys, interviews, and document analysis and examined to provide context for the information that had been gathered.

2.5 Method of Data Analysis

The initial portion of the questionnaire, which pertained to biographical information, was analyzed and interpreted through the use of descriptive techniques, specifically percentages. Additionally, the data collected via closed-ended questions was encoded and analyzed statistically using SPSS version 23.

To evaluate the differences between two indepen-

dent groups of data collected from schools A and B, an independent t-test was utilized. This statistical method is appropriate for comparing the means of two sets of data (Duffy & Orlandi, 2008; Watson & Moritz, 1998). Thematic analysis and semi-structured interviews were employed to obtain words, phrases, statements, and paragraphs to narrate the qualitative data. The qualitative data were utilized as a complement to the quantitative data.

3 Results and Discussion

3.1 Practice of Teachers’ Performance Appraisal

Accordingly, Tables One, Two, and Three, followed by a data description, present the practice that was measured in terms of accessibility, frequency of occurrence, and appraisers in a teacher’s performance appraisal in the targeted school.

Respondents were requested to indicate whether a formal performance appraisal (PA) system existed in their schools. The results revealed that 61 teachers (98.4%) affirmed the presence of an established PA practice within their institutions. In contrast, only one respondent reported that no

formal performance appraisal system was in place.

Concerning the frequency of occurrence displayed in the above, 50 (80.6%) respondents confirmed that performance appraisal activities were held in their school semiannually or twice a year. The remaining 8, 3, and 1 teachers assured that they were appraised once a year, three times a year, and four times or more a year.

Regarding appraisers, a majority of the respondents (48, 77.4%) claimed that directors or vice directors; the rest (six, 1, 4, and 0 respondents) discovered that department heads, student council members, members of PTA, cluster supervisors, and all were involved in undertaking teachers’ PA, respectively.

In the same mood, interviewees asserted that there is a formal appraisal practice in their school to achieve different goals. Interviewee TR revealed that "performance appraisal in our school aims to enhance better learning and achieve the school’s intended outcome". Similarly, interviewee WMH explained that "the aim of PA in the school is to identify the strengths and weaknesses of teachers and coach them on their limitations to help my school achieve its intended purpose".

Table 1: Availability of Teachers’ Performance Appraisal

Is there formal appraisal process in your school?						Total	
Yes	%	No.	%	I have no idea	%	No.	%
61	98.4	1	1.6	0	0	62	100

Table 2: Frequency of occurrence of Teachers’ Performance Appraisal

How often is performance appraisal conducted in your organization? Per year								Total	
Once	%	Two times	%	Three times	%	Four times and above	%	No.	%
8	12.9	50	80.6	3	4.8	1	1.6	62	100

Table 3: About appraisers of Teachers’ Performance Appraisal

In your school, who are involved in the process of teachers’ performance appraisal?											
Directors &/or Vice Directors	%	Department heads	%	Student	%	PTA	%	Supervisors	%	All	%
48	77.4	6	9.7	3	4.8	1	1.6	4	6.5	0	0

Table 4: Independent samples *t* – *test* for carry out of performance appraisal

	Sig.	Sig. (2-tailed)
Is there a formal appraisal system in your school?	.255	.576 .323
How often is performance appraisal conducted in your school?	.829	.666 .685
Who are involved in the process of teachers' performance appraisal?	.019	.233 .355

Note: The Sig (2-tailed) is a two-tailed p-value used to evaluate the null hypothesis that the mean is not equal to 50, and a value below the alpha level indicates statistically significant difference from zero.

Likewise, interviewee TH asserted that "the purpose of PA in my school is to motivate best-performing teachers and punish unsuccessful teachers, provided that both accountability and motivational purposes are materialized". Similarly, interviewees SUP and SC explain, "The results of our twice-a-year teacher performance evaluation at our school will be used as a measure of progress and other benefits". This tells the researcher that performance is being assessed in the schools. The independent *t* test results in terms of availability of formal PA in each school, regarding the frequency of PA in the two schools, and concerning who is involved in the process of teacher PA ($D = 0.576$ and p - value > 0.05), ($D = 0.666$ and p - value > 0.05), and ($D = 0.355$ and p - value > 0.05) respectively, show that the mean of the appraisal practice didn't bring any statistically significant variation between schools, implying that the accuracy of the performance appraisal practice was almost the same across schools.

3.2 Purpose of Teacher's Performance Appraisal

The purpose of a teacher's performance appraisal is to assess their effectiveness and provide feedback on their teaching methods and strategies.

Regarding the purpose of the performance appraisal, 13 items of questions were arranged and disseminated to teacher respondents. Of these questions, the first five focus on the purpose of PA being to improve instruction; the second three items of questions focus on teachers' professional development; the third two items of questions weigh on aiding school administrators in making decisions; and the last three items of questions concentrate on motivating teachers to take more responsibility.

Therefore, evidence respondents stated that PA experience in their respective schools was successful in improving instruction (mean = 4.51), enhancing teachers' professional development (mean = 4.57), aiding school administrators in passing decisions (mean = 4.17), and motivating teachers for more responsibility (mean = 4.23). This shows that a significant number of participants had the view that their PA had met its intended purpose.

The interviewees asserted that the purpose of PA is theoretically to achieve different goals. Nevertheless, the data from Interviewee TR didn't verify the findings presented in the table above. They brightly explained that the PA purpose in their respective schools is not meant to secure periodic promotion, motivation, and scholarship. Likewise, concerning improving classroom preparation, however, they clarified that it had made an insignificant contribution. According to Interviewee SUP, "Currently, teachers' performance appraisal is used to upgrade the teacher's education status; otherwise, the contribution of performance appraisal in rewarding outstanding teachers in terms of finance, certificates, and moral support is negligible".

In the same vein, interviewees WMH and TH make clear that the purpose of PA is right now for "paper value", implying that PA experience in the sample schools rarely contributed to improving instruction, enhancing teachers' professional development, or making valid administrative and motivational decisions.

Independent *t*-test results under the subject of improving instruction ($D = 0.473$ and p value > 0.05) indicate that there is no significant difference between the two schools. However, on the topic

Table 5: Performance appraisal purpose

No	Items related purpose of performance appraisal		1	2	3	4	5	Total	Mean
1	To improve the teaching-learning process	F	-	-	3	12	47	62	4.71
		%			4.8	19.4	75.8	100	
2	To identify the strengths and weaknesses of the teacher	F	-v	1	5	15	41	62	4.55
		%	-	1.6	8.1	24.2	66.1	100	
3	To decide on teacher's salary improvement	F	4	1	11	11	35	62	4.16
		%	6.5	1.6	17.7	17.7	56.5	100	
4	To enhance students' academic achievement	F	-	1	3	18	40	62	4.56
		%	-	1.6	4.8	29.0	64.5	100	
5	To offer a high-quality education	F	-	-	7	12	43	62	4.58
		%	-	-	11.3	19.4	69.4	100	
6	To identify training needs (training purpose)	F	-	5	4	10	43	62	4.47
		%	-	8.1	6.5	16.1	69.4	100	
7	To link teachers needs with the organizational goal	F	1	2	9	24	25	62	4.79
		%	1.6	3.2	14.5	38.7	40.3	100	
8	To help teachers develop professionally	F	-	1	8	15	38	62	4.45
		%		1.6	12.9	24.2	61.3	100	
9	To provide information on teachers' promotion & transfer	F	2	4	8	19	29	62	4.11
		%	3.2	6.5	12.9	30.6	46.8	100	
10	To control the overall performance of teachers	F	1	2	9	20	30	62	4.23
		%	1.6	3.2	14.5	32.3	48.4	100	
11	To improve the motivation of teachers	F	-	2	6	17	37	62	4.44
		%	-	3.2	9.7	27.4	59.7	100	
12	To provide feedback on teachers' performance	F	1	2	5	15	39	62	4.44
		%	1.6	3.2	8.1	24.2	62.9	100	
13	To serve as the basis for reward and punishment	F	6	2	11	20	22	62	3.82
		%	9.7	3.2	17.7	32.3	35.5	100	

The values are; strongly disagree =1; disagree =2; Undecided =3; agree =4; strongly agree = 5

of identifying the strengths and weaknesses of teachers ($D = 0.038$ and P value 0.05), this shows that there is a significant difference between the two schools.

Regarding performance appraisal criteria Performance criteria are statements of standards used for measuring job-related performance. For performance to be effective, it should encourage the participation of employees in designing the organizational goal up to its implementation. This is also true in the education system, where all stakeholders have to participate in all aspects of educational development activities.

As illustrated from Table 6, items 1–5, in the teachers' appraisal criteria, the group of respondents

responded with their disagreement with the participation of teachers in the formulation of TPA, the clarity of the criteria objectives, the relevance of the TPA criteria, the measurements of teachers' professional competency, and the appraisal criteria that measure teachers' motivation to work.

Almost all respondents showed their disagreement, and the mean values were less than the expected mean, which was 2.44. This implies that the entire group of respondents opposed the participation of teachers in the formulation of the TPA. However, Q2, 3, 4, and 5, with respective mean values of 4.03, 4.29, 4.02, and 4.24, indicated that the group of respondents agreed on the issues related to performance appraisal criteria.

Table 6: Regarding Performance appraisal criteria

No	Items related to criteria of performance appraisal		1	2	3	4	5	Total	Mean
1	There is the participation of teachers in the formulation of criteria	F	6	6	9	21	20	62	2.44
		%	9.7	9.7	14.5	33.9	32.3	100	
2	The criteria used are clear	F	2	4	11	18	27	62	4.03
		%	3.2	6.5	17.7	29	43.5	100	
3	The criteria used are relevant to the purpose of PA	F	2	2	6	18	34	62	4.29
		%	3.2	3.2	9.7	29	54.8	100	
4	The appraisal criteria objectively measure teachers' professional competence	F	2	5	10	18	27	62	4.02
		%	3.2	8.1	16.1	29	43.5	100	
5	The appraisal criteria measure teachers' motivation to work	F	2	2	6	21	31	62	4.24
		%	3.2	3.2	9.7	33.9	50	100	

The values are; strongly disagree =1; disagree =2; Undecided =3; agree =4; strongly agree = 5

Table 7: Independent samples T-Test for performance Appraisal criteria

Items	Sig.	Sig. (2-tailed)
There is a high participation of teachers in the formulation of criteria	.101	.216
		.276
The criteria used are clear	.232	.135
		.185
The criteria used are relevant to the purpose of TPA	.471	.917
		.926
The appraisal criteria objectively measure teachers' professional competence	.721	.949
		.950
The appraisal criteria measure teachers' motivation to work	.504	.852
		.867

Note: The Sig (2-tailed) is a two-tailed p-value used to evaluate the null hypothesis that the mean is not equal to 50, and a value below the alpha level indicates a statistically significant difference from zero.

The interviewee believed that the criteria were already developed by the Ministry of Education framework and adopted by the SNNPR education bureau. Speak with SUP and TH. It is important to note that teachers do not participate directly in the preparation process; they know what is expected of them when they enter into a goal agreement with the principal before starting work. SC has no information regarding a teacher's participation in the formulation of PA criteria. The independent t-test result regarding performance appraisal criteria, items of questions related to the participation of teachers in the formulation of criteria, the clarity of criteria used, the criteria relevant to the purpose of PA, the appraisal criteria measurement, and teachers' professional competence, and the appraisal criteria measure teachers' motivation to work.

Regarding Performance appraisal problems/ challenges

Since teachers' "performance appraisal is a complex activity, schools have faced problems in practicing it". These problems may be related to the skill and competence of evaluators; the operational process of the appraisal scheme and the perception of teachers about performance appraisal; the bias of evaluators; the maturity level of evaluators; and the criteria. In this regard, some (nine) of teachers' performance appraisal problems were listed, and 62 of them "were asked to rate those problems based on the Likert rating scale: very highly serious (5), highly serious (4), moderately serious" (3), minimally serious (2), and very minimally serious (1). According to Asfaw (2021), the mean scores from the data analysis were interpreted as 0.05–1.49 (very low), 1.5–2.49 (low), 2.5–3.49 (medium), 3.5–4.49 (high), and above 4.5 (very high).

Table 8: Regarding Performance appraisal problems/challenges

No	Items related to problem and challenges of PA		1	2	3	4	5	Total	Mean
1	Appraisers lack of the necessary knowledge & skill	F	9	8	13	17	15	62	3.34
		%	14.5	12.9	21	27.4	24.2	100	
2	Appraisers lack of the necessary experience	F	9	8	19	12	14	62	3.23
		%	14.5	12.9	30.6	19.4	22.6	100	
3	Poor administration of the overall process	F	10	9	15	15	13	62	3.19
		%	16.1	14.5	24.2	24.2	21	100	
4	Shortage of pre-appraisal discussion	F	9	10	11	13	18	62	3.5
		%	14.5	16.1	17.7	21	29	100	
5	Lack of post-appraisal discussion	F	2	4	8	19	29	62	4.11
		%	3.2	6.5	12.9	30.6	46.8	100	
6	The presence of negative perception of teachers about PA	F	6	7	12	19	18	62	3.58
		%	9.7	11.3	19.4	30.6	29	100	
7	The presence of inadequate and inappropriate appraisal criteria	F	10	6	17	17	12	62	3.24
		%	16.1	9.7	27.4	27.4	19.4	100	
8	Bias of evaluators that can replace organizational standards by personal values in the evaluation process	F	9	13	12	12	16	62	3.21
		%	14.5	21	19.4	19.4	25.8	100	
9	Students aren't mature enough to properly evaluate their teachers	F	8	7	9	15	23	62	3.61
		%	12.9	11.3	14.5	24.2	37.1	100	

In response to items 1–9, in the Performance Appraisal Problems/Challenges, the teacher respondents reply with the necessary knowledge and skills, the necessary experience, administration of the overall process, pre-appraisal discussion, and post-appraisal discussion, the perception of teachers about PA, inappropriate appraisal criteria, the bias of evaluators, and maturity to properly evaluate their teachers. Almost all respondents showed their agreement, and the mean values were greater than the expected mean (3:00), which were 3.34, 3.23, 3.19, 3.5, 4.11, 3.58, 3.24, 3.21, and 3.61.

On the other hand, the responses obtained from interviewees TR and WMH were forwarded with the following suggestions: "Due to the lack of knowledge, skill, and experience of appraisers, they found that managing the appraisal process in their school was one of the more difficult tasks". Interviewee SUP also argued that "post-appraisal discussion is the most serious challenge in the teacher performance appraisal process". Interviewees TH and SC, on their part, forwarded the following as the major constraints: "Mostly higher levels of academic achievement" students are involved in the appraisal of teachers' performance. According to them, students ranked 1-3 mostly participated in teacher evaluations. Students do not have any training in PA, and sometimes they use the appraisal

as a revenge mechanism. Most students were not aware of the purpose of TPA, so they simply filled out the form. Some students are regularly "biased in marking on matters unrelated to the performance of teachers".

4 Discussions

4.1 Practice of Teachers' Performance Appraisal

Regarding availability, in those "sample schools, teachers' performance is evaluated on a regular basis or by procedure, and the schools have a planned way of letting the teachers know where they stand and how they are progressing". This means very large numbers of teachers were assured of the existence of performance appraisal practices in primary schools in the mentioned city.

Regarding the frequency of the occurrence of performance appraisals in the selected school, questionnaires and interviewee respondents revealed that there is teacher performance appraisal practice that occurs twice a year. Teacher evaluation should be conducted twice at the end of each semester (MOE, 1997). This is because not only are the rules established twice a year, but instructor performance ratings are also calculated each semester as part of the student assessment process. In this sit-

uation, teacher performance evaluation, formative use of performance appraisal, or making the best use of teacher evaluation data poses several issues and is used to decide on career promotion, give out performance advantages, or penalize instructors who don't perform at the same level.

Regarding appraisers of TPA, the result shows that the majority of participants (48, or 77.4%) claimed that the school principals are the main ones, and it seems like a solitary agent who evaluates their performance through inbuilt or other parts of performance related to the school system. However, appraisal of the performance of teachers requires participation from different stakeholders (MOE, 2002). Nonetheless, this shows that either the school leader is not willing to use other possibilities, such as self-appraisal, colleagues, etc., or may have other justifications. This issue needs further investigation. If employees are not allowed to evaluate themselves, they will become highly defensive during the appraisal review and may refuse to accept the evaluation result. It is also critical to understand the weak points of your coworkers.

The independent t test results in terms of the availability of formal PA in each school, regarding the frequency of PA in the two schools, and who is involved in the process of teaching PA ($D = 0.576$ and $p\text{-value} > 0.05$), ($D = 0.666$ and $p\text{-value} > 0.05$), and ($D = 0.355$ and $p\text{-value} > 0.05$), respectively, show that the mean of the appraisal practice didn't bring any statistically significant variation between schools, implying that the accuracy of the performance appraisal practice was almost the same across schools.

4.2 Performance appraisal purpose

In goal theory, Locke and Latham (2019) highlight mechanisms that connect goals to performance outcomes. The other theory Control theory emanates from the work of Emile Durkheim and focuses attention on feedback as a means of shaping behavior (Glad & Ljung, 2018). The social cognitive theory was developed by Bandura. According to Beauchamp *et al.* (2019), it is based on his central concept of self-efficacy. This suggests that what people believe they can or cannot do has a powerful impact on their performance.

Regarding the purpose of performance appraisal, the study results indicated that teachers are not competing with each other based on their PA results for promotion or demotion. The investigation made it clear that the teachers' outcomes were undesirable. However, the wrong attitude of carelessly "using it" as if teachers must be promoted; there is also concrete quota promotion (professional training, upgrading) from teachers in the school. The purpose of performance appraisal was misused or interpreted negatively in this case.

In line with this, it is possible to explain why the interviewees reflect the purpose of PA destructively. The researcher suggested that the current system of performance appraisal is conducted only for procedure purposes, and there is no way to accomplish its purpose properly. When there is an opportunity for benefit, the quota and the type of subject will come together, and the teacher who does not perform well will possibly benefit from the quota. In the same way, promotion is associated with a pay raise, but all teachers benefit from promotion, even if they do not perform well. This means that the purpose of teacher evaluation has failed.

Independent t test results under the subject of improving instruction ($D = 0.473$ and $p\text{ value} > 0.05$) indicate that there is no significant difference between the two schools. However, on the topic of identifying the strengths and weaknesses of teachers ($D = 0.038$ and $P\text{ value } 0.05$), this shows that there is a significant difference between the two schools. Therefore, the aim of performance appraisal in school B is to help the appraisers identify the strengths and weaknesses of teachers and support them based on evidence accordingly, but in school A, the topic of identifying the strengths and weaknesses of teachers is not reasoned. This indicates that in school B, there is formative assessment practice, while in school A, there is no formative assessment practice at all. Glad & Ljung (2018) and Ozbay (2019), a control theory, focus attention on feedback as a means of shaping behavior. Thus far, the rest of the questions Q 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, and 20 have independent t test results ($D = 0.473, 0.724, 0.265, 0.584, 0.996, 0.772, 0.086, 0.846, 0.42, \text{ and } 0.588$, respectively). This shows that with a $P\text{-value} > 0.05$, there is

no significant difference between the two schools regarding the above points of the question of TPA.

4.3 Performance appraisal criteria

In this context, participants were requested to assess the appraisal criteria currently implemented in the sampled schools. The findings indicate that the calculated mean scores (2.44, 4.03, 4.29, 4.02, and 4.24), together with the overall weighted mean of 3.8, exceeded the benchmark or ideal mean value of 3. However, the specific item with a mean score of 2.44 suggests that most respondents expressed disagreement with that particular statement. As emphasized by Nadot (2022), even minor procedural shortcomings can undermine the effectiveness of an appraisal system and potentially result in its failure. Consistent with this concern, respondents indicated dissatisfaction regarding teachers' limited involvement in the development and formulation of performance appraisal criteria.

In line with this, the interview group made public that there was no teachers' participation in the formulation of the TPA criteria. The interviewee believed that the criteria were already developed by the Ministry of Education Framework and adopted by the SNNPR Education Bureau. The interview group made it clear that teachers do not participate directly in the preparation process; they know what is expected of them when they enter into a goal agreement with the principal before starting work. It is also important to understand if there are any questions or concerns that you do not understand. According to the results of the survey and the interview, one of the reasons why teachers do not participate directly in the preparation of the assessment criteria is that the assessment criteria have already been developed by the Ministry of Education and sent to all regions. The second is that when teachers agree on a goal with the principal, they discuss the key, major, and minor tasks in depth and ensure understanding.

According to the information obtained from the document, the criteria for the evaluation of the performance of teachers transferred from the Ministry of Education and the criteria prepared by the principals of both sampled schools are the same. This does not make teachers see themselves as stake-

holders who attempt to achieve the school's goals and objectives rather than being directly involved in setting standards. As Sewagegn (2019) stated, when preparing the teacher's performance appraisal standard, if the teacher knows as much as possible about the standard that is being evaluated, there will be many benefits.

The independent t-test result regarding performance appraisal criteria, items of questions related to the participation of teachers in the formulation of criteria, the clarity of criteria used, the criteria relevant to the purpose of PA, the appraisal criteria measurement and teachers' professional competence, and the appraisal criteria measure teachers' motivation to work (p value > 0.05) shows that the mean of the appraisal exercise regarding the criteria didn't bring any statistically significant variation between schools, implying that the accuracy of the performance appraisal criteria was almost the same in two schools.

4.4 Performance appraisal problems and challenges

In this regard, nine of the teachers' performance appraisal "problems were listed, and 62 teachers were asked to rate those problems based on the Likert rating scale very highly serious (5), highly serious (4), moderately serious (3), minimally serious (2), and very minimally serious (1)". According to Asfaw (2021), the mean scores from the data analysis were interpreted as 0.05–1.49 (very low), 1.5–2.49 (low), 2.5–3.49 (medium), 3.5–4.49 (high), and above 4.5 (very high).

Therefore, the mean score of participants in the three items, Q.36, 38, 41, and 30, was related to the maturity level of evaluators (3.61), the negative perception of teachers towards performance appraisal (3.58), the shortage of pre-appraisal discussion (3.5), and the shortage of post-appraisal discussion (4.11).

Participants were rated as having highly serious

The findings further indicate that several issues were identified as serious concerns, with mean scores ranging between 3.5 and 4.49. In contrast, the mean ratings for items Q33, Q34, Q35,

Q39, and Q40—addressing weak administrative support (3.19), insufficient skill and knowledge (3.34), limited experience in conducting performance appraisal (3.23), inadequacy of appraisal criteria (3.24), and evaluator bias (3.21)—fell within the range of 2.5 to 3.49. These values suggest that respondents perceived these factors as moderately serious challenges affecting the appraisal process.

Similarly, interview data revealed that deficiencies in appraisers' knowledge, competence, and practical experience made the management of the appraisal system difficult within the schools. Participants emphasized that inadequate professional training for evaluators and minimal teacher involvement during the development of performance appraisal criteria constituted the most critical challenges facing teacher performance appraisal (TPA) practices. Moreover, it was noted that students—often those with relatively high academic achievement—participate in evaluating teachers' performance (Almutairi & Shraid, 2021). However, students typically lack formal preparation in performance appraisal procedures, and in some instances, the evaluation process may be influenced by personal motives, including the use of appraisal as a form of retaliation.

As LEMJI (2019) also argued, post-appraisal discussions between the assessed and the appraiser are a highly serious challenge in the teacher performance appraisal process. The t test analysis for the data in both (school A and school B) government primary schools indicates that there is no significant difference between the two schools.

5 Conclusions

Formal teachers' performance appraisals are often conducted twice a year. The practice is measured in terms of accessibility and frequency of occurrence. Although the guidelines state that many stakeholders are involved, the principal is the one who has the highest share. This is outside of the guidelines and can lead to biased results. Every teacher gets promoted when it comes to performance; every teacher gets a pay raise and benefits. On the contrary, teachers who do not show the proper quality and results in the work of the school may be prevented from progressing under the excuse of quotas.

Either way, the process is unfair. Teachers are not competing with each other and are promoted as long as they meet the criteria. Teachers expressed satisfaction with the suitability of the standards. However, they feel that other points that should be included have been left out because they were not involved in developing the appraisal criteria. Here, the evaluation criteria come from the Ministry of Education, but they remain localized. Rather than presenting another alternative, the principals copied the requirement from the guidelines and gave it to us to implement, so this opens the door to influence and can be a hindrance to achieving the school's goals.

According to the findings, insufficient training for evaluators is a critical issue in addressing the challenge and problem at hand. For example, the evaluators selected from among the students use the appraisal of the teachers to take revenge on the appraiser. Moreover, the lack of teacher participation in the development of performance evaluation criteria for teachers is a significant area for improvement. The absence of post-evaluation discussions, teachers' difficulty in identifying their weaknesses, and their low regard for the importance of school improvement are additional factors that contribute to the problem. Consequently, teachers may perceive evaluation results as lacking relevance to their professional growth.

6 Recommendations

Drawing from the aforementioned conclusions, the following recommendations are suggested: Schools in the study area must implement formal teacher performance appraisal systems and practices, typically conducted on a biannual basis. However, it is crucial to note that these practices often appear superficial and aim to fulfill bureaucratic requirements rather than foster genuine improvement. A fundamental shift in perspective is therefore necessary. To address this issue, school leaders ought to utilize teacher performance appraisals as a means of driving strategic, developmental, and communicative initiatives, with the ultimate goal of enhancing teacher academic and professional competencies as well as overall school performance.

Concerning the intended purpose of performance

appraisal, the author of the study proposes the following recommendations: school administrators should carefully consider the objective of teacher performance evaluations and carry them out according to relevant guidelines and objectives. Additionally, it is suggested that each school establish its own comprehensive teacher performance evaluation plan to optimize the system's operation and implementation. This plan, along with the school's strategic objectives, should be communicated to all stakeholders and experts. Furthermore, it is recommended that school principals design tailored training programs for teachers and implement a performance-based reward system to motivate superior performers.

The criteria for evaluating the performance of teachers are primarily formulated at the central level, and there is a sense of dissatisfaction among teachers that these criteria do not take into account the contextual situation, including differences between schools and the lack of participation of teachers in the process. To address this issue, the Education Bureau should ensure that assessors are involved in the preparation of evaluation criteria in appropriate circumstances. School leaders and experts should also consider the current situation of the schools and contextualize the appraisal criteria in a way that can enhance the capacity of teachers. To avoid problems with teacher performance evaluations, it is important to use interpersonal communication and set evaluation criteria without personal or political interference. In addition, recruiting older students to serve as student appraisers and providing adequate training on academic and social relationships can help address the issue.

Conflict of Interests

The authors declare that there is no conflict of interest.

References

- Almutairi, T. S. & Shraid, N. S. (2021). Teacher Evaluation by Different Internal Evaluators: Head of Departments, Teachers Themselves, Peers and Students. *International Journal of Evaluation and Research in Education*, 10(2):588-596.
- Asfaw, B. (2021). Teachers Performance Appraisal Practices in Government Schools of Addis Ababa. Doctoral Dissertation, St. Mary's University.
- Baye, A. (2021). Teachers' Performance Appraisal Practices in Government Schools of Addis Ababa.
- Beauchamp, M. R., Crawford, K. L., & Jackson, B. (2019). Social cognitive theory and physical activity: Mechanisms of behavior change, critique, and legacy. *Psychology of Sport and Exercise*, 42:110-117.
- Chen, J., Cheng, H., Zhao, D., Zhou, F., & Chen, Y. (2022). A Quantitative Study on the Impact of Working Environment on the Well-Being of Teachers in China's Private Colleges. *Scientific Reports*, 12(1):1-9.
- Creswell, J. W. & Zhang, W. (2009). The Application of Mixed Methods Designs to Trauma Research. *Journal of Traumatic Stress: Official Publication of the International Society For Traumatic Stress Studies*, 22(6):612-621.
- Dawadi, S., Shrestha, S., & Giri, R. A. (2021). Mixed-methods research: A discussion on its types, challenges, and criticisms. *Journal of Practical Studies in Education*, 2(2):25-36.
- Glad, T., & Ljung, L. (2018). Control theory. CRC press.
- Dhivyadeepa, E. (2015). Sampling techniques in educational research. Lulu. com.
- Duffy, A. & Orlandi, A. (2008). A review of statistical methods for comparing two data sets. *The Applied Computational Electromagnetics Society Journal (ACES)*, 90-97.
- Englander, M. (2012). The interview: Data collection in descriptive phenomenological human scientific research. *Journal of phenomenological psychology*, 43(1):13-35.
- Islami, X., Mulolli, E., & Mustafa, N. (2018). Using Management By Objectives As A Performance Appraisal Tool for Employee Satisfaction. *Future Business Journal*, 4(1):94-108.

- Kelly, S. E., Bourgeault, I., & Dingwall, R. (2010). Qualitative interviewing techniques a styles. *The SAGE handbook of qualitative methods in health research*, 19:307-326.
- Lemji, S. (2019). An Assessment of Practices and Challenges of Teachers' Performance Appraisal System Enhancing Organizational Performance in Secondary Schools of Gurage Zone (Doctoral dissertation).
- Locke, E. A. & Latham, G. P. (2019). The development of goal setting theory: A half century retrospective. *Motivation Science*, 5(2):93.
- Manna, R. & Mete, J. (2021). Population and Sample. *International Journal of Research and Analysis in Humanities*, 1(1):30-30.
- Ministry of Education (2002). Education Sector Development Program Action Plan-II (2001/2-2004/5). Addis Ababa: FDRE Ministry of Education of Ethiopia
- Mohammed, A. A. (2020). Review Paper on University Teachers Performance Appraisal. *Open Access Library Journal*, 7(7):1-18.
- Monari, D. G. & Wanjau, K. L. (2022). Factors Affecting Utilization of Performance Appraisal Practices in the Telecommunications Industry: A Survey Of Telecommunication Firms In Kenya. In *Proceedings of the Sustainable Research and Innovation Conference* (Pp. 228-232).
- Mulualem, Y. G., Mulu, Y. E., & Gebremeskal, T. G. (2022). Relations of Previous English Achievement with English Learning Beliefs and College Academic Achievements. *HELIYON*, E09841.
- Nadot, Y. (2022). Fatigue from Defect: Influence of Size, Type, Position, Morphology and Loading. *International Journal of Fatigue*, 154, 106531.
- Negash, E., Asmare, A., & Eshete, S. K. (2021). Effects of Employee Motivation on Organizational Performance at Ethiopian Telecom South West Region Jimma. *Ijhcmm (International Journal of Human Capital Management)*, 5(1):28-41.
- Ozbay, H. (2019). *Introduction to feedback control theory*. CrC Press.
- Ozkeser, B. (2019). Impact of training on employee motivation in human resources Management. *Procedia Computer Science*, 158:802-810.
- Sewagegn, A. A. (2019). A Study on the Assessment Methods and Experiences of Teachers at an Ethiopian University. *International Journal of Instruction*, 12(2):605-622.
- Shugute, E. (2020). Practices and Challenges of Managing Teachers 'performance Appraisal in Secondary Schools of Kembatta Tembaro Zone (Doctoral Dissertation).
- Vansteenkiste, M., Ryan, R. M., & Soenens, B. (2020). Basic psychological need theory: Advancements, critical themes, and future directions. *Motivation and emotion*, 44(1):1-31.
- Watson, J. M. & Moritz, J. B. (1998). The beginning of statistical inference: Comparing two data sets. *Educational Studies in Mathematics*, 37(2):145-168.
- Zeng, Z. & Qi, L. (2021). "Internet+ Artificial Intelligence" Human Resource Information.
- Nadot, Y. (2022). Fatigue from Defect: Influence of Size, Type, Position, Morphology and Loading. *International Journal of Fatigue*, 154, 106531.
- Nadot, Y. (2022). Fatigue from Defect: Influence of Size, Type, Position, Morphology and Loading. *International Journal of Fatigue*, 154, 106531.