Quest Journals
Journal of Research in Humanities and Social Science
Volume 14 ~ Issue 11 (November 2026) pp: 05-13
ISSN(Online):2321-9467
www.questjournals.org

Research Paper

Teacher and Peer Feedback in EFL Speaking: A Critical Narrative Review of Oral Performance, Learner Perceptions, and Feedback Quality

Shaila Kajol Munmun
[1] *(Department of English, City University, Bangladesh)*

ABSTRACT
This critical narrative review intended to be a critical examination of the relationships between teacher and peer feedback and speaking outcomes in English as a Foreign Language (EFL) learning, focusing on learners' perceptions and feedback quality. In October 2026, a source-led search and verification of scholarly publications was conducted, but not a systematic search of the databases. Seven empirical studies in EFL and two additional studies in oral-presentation feedback in non-EFL higher education were chosen for close comparison. As a secondary context, one systematic review of mobile-assisted peer feedback was taken up. In the studies included in this sample, both sources of feedback were equally effective, or were ineffective in some instances. In some studies of feedback quality and presentation behavior, the teacher feedback showed advantages, while some EFL studies indicated that higher performance levels benefited from peer correction. Other comparisons revealed gains under both conditions. The advantages of teacher feedback were found in some studies of the feedback quality and presentation behavior, and some of the EFL studies found higher performance level benefited from peer correction while other studies found gains under both conditions. The links between self-efficacy and speaking performance were found to be somewhat indirect and the preferences were not always congruent with the outcomes assessed. Results on inter-rater reliability and interpersonal bias also indicate distinct concerns with perceived trust, quality of comments and score consistency. The review suggests that feedback needs to be evaluated based on its appropriateness for the task to be spoken, the assessment procedures, as well as the opportunities provided for learners to interpret and respond to comments. It points out a need of studies focused on the assessment of performance, perceptions of the learner, and feedback quality in the same design.

KEYWORDS: *EFL speaking; teacher feedback; peer feedback; oral performance; learner perceptions; feedback quality; speaking self-efficacy*

Received 25 Oct., 2026; Revised 04 Nov., 2026; Accepted 06 Nov., 2026 © The author(s) 2026.
P*ublished with open access at* ***www.questjournals.org***

I. INTRODUCTION

The process of speaking in English as a Foreign Language (EFL) classrooms cannot be developed without feedback. In communicating with others, learners have to organize ideas, ideas, pronunciation, grammar and lexical options, fluency and coordinate a combination of these factors together. Constructive feedback can also help focus on areas of an oral performance that should be improved. It is not the provider of this that makes it have educational value. The comments should be pertinent to the task, clear to the learners, and detailed enough to guide the future practice.

The feedback of the teacher and peers occurs in various positions. Educators usually introduce more linguistic and evaluation knowledge and are able to relate their remarks to teaching objectives. However, there might be constraints on teacher remarks due to the time of classroom usage, and the effectiveness of the information might be limited by the ability of learners to take action. Peer feedback provides students with an active experience in assessing oral performance and understanding assessment criteria. It can extend the possibilities of feedback in a speaking lesson, though students may also come up with vague remarks, make wrong identification of errors or be reluctant to criticize classmates. This is not only a comparison between a qualified assessor and an unqualified one; it involves as well the purpose and design of the feedback activity.

Research has indicated varying results. Au and Bardakci (2020) reported better oral performance for 33 EFL students of a Turkish university in both teacher and peer feedback conditions, but a more significant improvement in the teacher-feedback group. Mahvelati (2021) has found a stronger speaking development when using peer corrective feedback in the context of Iranian EFL learners, although the learners expressed worries about the validity and social implications of peer correction. A study by Li and Hu (2024) also revealed the benefits of peer correction in speaking accuracy and fluency in a Chinese university. Thai oral-presentation comparison by Wanchid and Charoensuk (2024) did not focus on a binary choice but on three types of feedback arrangements, which demonstrated the significance of the delivery mode and task design.

The measures of the results in this literature cannot be interchanged. Speaking tests reveal elements of overt performance; speaking self-efficacy deals with beliefs of the learners about their ability to accomplish speaking tasks; perceptions are how the participants perceive feedback. Scoring reliability is an agreement or consistency in assessment, whereas feedback quality is properties of the comments themselves. A seemingly good result in one area cannot be used to create a benefit in another. As an example, Au and Bardakcizadeh (2020) did not observe statistically significant differences in measured oral performance and speaking self-efficacy in their sample. Li et al. (2022) focused on the reliability of peer scores and utility of comments, Tanaka (2025) conducted an analysis of peer rating bias based on friendship. These works shed some light on various aspects of the feedback process.

There are some closely related studies that do not involve EFL instruction. In two studies by van Ginkel et al. (2017a, 2017b), oral-presentation courses in Dutch higher education, as opposed to EFL speaking courses, were used. They are very helpful in the above comparison evidence of the feedback processes and presentation behavior but they cannot be enumerated as the evidence itself regarding EFL language learning. This difference is important since the presentations in general higher education and EFL speaking activities can be based on similar performance skills without the need to test the same linguistic performance.

Weng et al. (2025) conducted a systematic review of 15 studies on mobile-assisted peer feedback of EFL speaking. It is not as technologically oriented as the current paper which takes into consideration teacher-peer comparisons, learner reactions and issues of assessment in the context of feedback schemes. This review is not intended to argue that teacher and peer feedback have hardly been compared. It is to investigate the alterations in the conclusions under different circumstances when separate outcomes and circumstances are maintained.

The review answers three questions: (1) What are comparative studies telling us about teacher and peer feedback in EFL speaking performance? (2) What is possible to determine regarding the perceptions of learners and, where measured directly, speaking self-efficacy? (3) What is the quality of feedback, assessment reliability, interpersonal, and study design that qualify those findings? These questions are used in the discussion to make an evidence-sensitive interpretation as opposed to a broad ranking of the two sources.

II. METHODOLOGY

2.1 Review design

A focused critical narrative review was used to compare a heterogeneous set of studies concerned with feedback on spoken performance. The source publications differ in task, population, intervention, measurement, and purpose. A narrative synthesis permits comparisons among these differences without implying that a pooled effect estimate or exhaustive systematic review has been produced. No original participant data were collected for this paper.

2.2 Source identification and checking

During the development of the manuscript, searches with specific keywords in the ERIC, publisher and journal websites, PubMed Central, and web-indexed scholarly records were conducted to find relevant publications, and bibliographic and factual verification was last updated on 23 October 2026. The search terms were related to teacher feedback, peer feedback, EFL speaking, oral presentation, speaking performance, speaking self-efficacy, and corrective feedback, peer assessment, rating reliability, trust, and learner perceptions. The search was focused on English-language 2015-2026 mostly. Relevancy of records was checked and publication details and findings were confirmed by having access to journal abstracts or full text. It was a source-based, non-exhaustive process: the review is not purported to have performed a searchable database-wide search, a PRISMA screening process, or a final number of records searched and excluded.

The sample of the studies was chosen in a purposive manner to answer the review questions. Research papers were included in the study that directly compared teacher and peer feedback in EFL speaking or provided

an explicit and relevant view of peer assessment of spoken English. Two non-EFL tertiary studies were kept as expressly delineated cross-context reports of the feedback quality and presentation competence. Empirical evidence was not based on studies that only involve feedback on writing and research which lacked a significant oral-performance or oral-assessment aspect. The corresponding review of mobile-assisted peer feedback was not considered as a separate primary study, but as secondary contextual literature.

2.3 Evidence base and classification

Nine empirical articles constitute the selected evidence base (Table 1). Five are EFL studies directly comparing feedback or assessment sources (Au & Bardakçı, 2020; Li et al., 2022; Li & Hu, 2024; Mahvelati, 2021; Wanchid & Charoensuk, 2024). Two additional EFL studies provide context on perceptions and peer-rating bias without testing a teacher-versus-peer feedback intervention (Charoensuk & Wanchid, 2025; Tanaka, 2025). Two Dutch higher-education presentation studies are included only as cross-context support (van Ginkel et al., 2017a, 2017b). The ten-item reference list also includes Weng et al. (2025), a systematic review used for contextualization. Accordingly, seven selected empirical articles involve EFL and two concern general higher-education oral-presentation work.

The word "comparison" is applied cautiously. Studies comparing corrective comments and subsequent language development are not equivalent to studies comparing ratings of recorded speeches. Likewise, perception studies and studies of bias do not independently test whether feedback improves speaking performance. Table 1 records both the useful finding and the main interpretive boundary of each article.

Table 1. Comparative evidence matrix for the nine selected empirical studies

Study and context	Evidence role / design	Finding relevant to this review	Interpretive boundary
Au & Bardakçı (2020); Türkiye; EFL	Direct comparison; 33 preparatory-school learners; speaking assessments and self-efficacy scale	Both feedback groups improved; teacher group had greater reported gains. No significant self-efficacy–performance correlation.	Small sample; outcome concerns speaking performance and self-efficacy, not feedback credibility.
Mahvelati (2021); Iran; EFL	Direct corrective-feedback comparison; speaking performance and learner perceptions	Peer-correction group improved more despite negative attitudes and concerns about trust, grades, and face.	Findings concern a specific corrective-feedback setting; perceptions and performance are distinct.
Li et al. (2022); China; EFL	Peer-versus-teacher assessment of 34 speech samples; 30 student and 8 teacher assessors	Two peers' holistic rating reliability approximated one teacher assessor; qualitative comments were similarly effective in the study.	Assessor reliability/comment evaluation, not evidence that a feedback intervention caused speaking gains.
Li & Hu (2024); China; EFL	Factorial study in six intact classes; peer/teacher/no correction and instructional mode	Peer correction outperformed teacher correction on accuracy, fluency, and academic passion; correction-by-mode interaction not significant.	Academic passion is not speaking self-efficacy; setting limits generalization.
Wanchid & Charoensuk (2024); Thailand; EFL	108 engineering students; oral teacher, oral peer, online written peer feedback	Feedback type and English achievement were associated with presentation outcomes; no significant interaction; positive group perceptions.	Peer conditions also involved teacher input; a pure feedback-source effect cannot be isolated.
Charoensuk & Wanchid (2025); Thailand; EFL	Contextual perception study; 374 students and 10 instructors	Both groups generally favored peer-feedback activities but differed on perceived effort and enjoyment.	Attitude evidence; does not directly compare learning gains under teacher and peer feedback.
Tanaka (2025); Japan; EFL	Contextual peer-rating study of prepared oral presentations	Modest friendship-related leniency and rater-gender differences in peer assessment.	Presenters did not receive peer ratings/comments; findings concern scoring bias, not feedback uptake.
van Ginkel et al. (2017a); Netherlands; non-EFL university	Cross-context: 95 filmed post-presentation feedback episodes; teacher, peers, tutor-guided peers	Teachers met more feedback-quality criteria; tutor-guided peers exceeded unguided peers on four criteria.	General university presentation training, not EFL; analyses feedback quality, not EFL language gains.

Study and context	Evidence role / design	Finding relevant to this review	Interpretive boundary
van Ginkel et al. (2017b); Netherlands; non-EFL university	Cross-context: quasi-experimental study of 144 first-year students	Teacher feedback most beneficial for presentation behavior; cognition and attitudes improved across sources.	General university speaking competence; limited transfer to EFL linguistic outcomes.

2.4 Approach to synthesis

The nine articles were compared in terms of setting, study purpose, feedback arrangements, outcomes, reported findings, and limits to interpretation. The analysis distinguished (a) speaking or presentation performance from feedback quality, (b) learners' perceptions from task-specific speaking self-efficacy, and (c) peer-rating reliability from interpersonal credibility. Findings were grouped by these distinctions and then compared across EFL and cross-context evidence. No effect-size aggregation, independent reanalysis of statistical results, or formal risk-of-bias scoring was carried out. The review's methodological limits are addressed together in Section 6 rather than repeated here.

III. LITERATURE REVIEW

3.1 Performance effects in EFL comparisons

The EFL comparison studies are not one-way. A total of 33 preparatory-school learners were followed by Au and Bardakçı (2020) in speaking assessment and found large and significant improvements in performance following teacher and peer feedback. The teacher feedback group had a greater gain on the final test. The comparison was carried out in a small group and this is only informative of that instructional context and should not be regarded as providing general evidence of teacher superiority.

The opposite has been stated in Iranian EFL speaking instruction by Mahvelati (2021) who found that peer corrective feedback led to more development than teacher corrective feedback. However, concerns about peer feedback, such as trust and grade or social consequences, were raised by learners. This is particularly relevant evidence due to the differences in the results of the performance and the choice of the participants. But, not all anxiety and not all reluctance is to be encouraged as a teaching measure – what is reported here is specific to the context reported.

Li & Hu (2024) compared the teacher, peer, and no-correction conditions in online and face-to-face classes. Peer correction achieved higher scores than teacher correction in the accuracy, fluency and academic passion of speaking in the university sample. The interaction between corrections by instruction was not significant. The study, therefore, lends support to the potential of peer correction in this design and not to the general proposition that online peer feedback is always "better" or "academic passion" can be replaced by speaking self-efficacy.

Wanchid and Charoensuk (2024) compared oral teacher feedback, oral peer feedback and online written peer feedback on English oral presentation given by 108 engineering undergraduates. They found that feedback type and English achievement level had an effect on performance, but did not find significant interaction between them. Pupils had positive views overall. A crucial criterion is that the peer conditions did not remove the role of teacher input, and the peer input is only represented in the nominal feedback source. The three conditions were also different in terms of their modality, making this study more useful to the discussion of feedback arrangements than a purely isolated teacher versus peer comparison.

Together, these four performance-based research studies illustrate the challenges of ordering the sources on a single performance scale. Correction of spoken errors is not the same as assessment following a prepared oral presentation. The skill level of learners, tasks, scope of comments and tracking progress is variable. Both results favorable and results favorable for peer feedback can be valid in their own design, as can results favorable and results favorable for teacher feedback.

3.2 Feedback quality and reliability: distinct kinds of evidence

A second set of findings relates to the properties of feedback or assessment, and not to learning gains. Thirty student assessors and eight experienced teachers were asked to assess 34 English speech samples by Li et al. (2022). They found that the reliability of two peer assessors may match that of one teacher assessor, with respect to the holistic score, when applied as per the procedures followed in their study. Peer qualitative comments and teacher qualitative comments were also found to be effective in their evaluation. This is evidence on the

possibility of peer assessment in a structured EFL speaking-rating task and should not be rephrased as a controlled intervention showing that peer feedback leads to an equivalent improvement in speaking.

Careful classification would be helpful in understanding why feedback design is important, as both of the van Ginkel studies contribute to this understanding. Van Ginkel et al. (2017a) did a content analysis on 95 recorded feedback interactions on 7 content and form criteria of feedback quality in a Dutch university. Teacher feedback was more consistent with the criteria and tutor guidance improved peer feedback on four criteria. Feedback was particularly effective on presentation behaviour in a related study on 144 undergraduates in which cognition and attitudes improved in all feedback conditions (van Ginkel et al., 2017b). Both studies refer to oral presentation teaching which is not part of an EFL learning intervention. They provide a plausible lesson plan on providing feedback to students and giving them guidance, rather than evidence on students' pronunciation, grammar or fluency development in English.

No necessary conflict between the scores teachers earned on a rubric for teachers' performance on a feedback quality dimension and the peer scores in Li et al. (2022). Not all of the studies are the same. The former examines the quality of feedback comments based on predetermined criteria; the latter looks at the reliability of holistic marks and usefulness of qualitative comments in an EFL speaking assessment format. Having a consistent score does not automatically mean that feedback is accurate, credible or followed by the receiver.

3.3 Speaking self-efficacy and perceptions

Speaking self-efficacy is a task-specific belief in ability to speak English, and is not a construct of motivation, enjoyment, general confidence, or a preference for a feedback source. The direct evidence of self-efficacy in the selected comparative studies mainly focuses on Au and Bardakçı (2020). They did not find a statistically significant correlation between speaking self-efficacy and oral performance in their 33 subjects. The non-significant result does not prove that there is no relationship in the broader EFL population, especially since this was a small sample. It does state that an increase in perceived ability is not necessarily an indicator of measured language development.

Evidence regarding learner perceptions is wider, but should not be interpreted as evidence of self-efficacy. Mahvelati (2021) indicated negative attitudes towards peer corrective feedback although there was greater improvement of the peer corrective feedback. Even though the outcomes of the presentation were different between the feedback conditions, the group perceptions were positive according to Wanchid and Charoensuk (2024). According to these studies, the preference and improvement as assessed may be different.

Charoensuk and Wanchid (2025) studied the perspectives of 374 learners and 10 teachers in an English oral-presentation class. In general, both groups were positive about the peer-feedback activities, and they had conflicting judgments regarding the amount of effort required and enjoyment. This is a perception-based study that identifies practical implementation issues, but does not measure comparative teacher/peer speaking gains. A view of perceptions in the chosen literature is that they can be considered information about acceptability and participants' experience rather than being a proxy for performance or feedback accuracy.

3.4 Interpersonal bias and the limits of perceived credibility

Feedback credibility is often talked about as being one variable. The studies chosen rather give indirect evidence on other separate properties, such as: trust in the person who made the comment, accuracy or specificity of comments, consistency of scores, and social influences on judgment. According to Mahvelati (2021), the participants in this study indicated trust and face concerns related to peer correction. While these experiences may affect whether or not a learner listens to advice, they are not evidence of whether or not a specific correction is linguistically correct.

Tanaka (2025) investigated peer evaluation of oral presentations prepared in English among Japanese EFL university students and found there were small effects of the friendship and rater-gender. This study is particularly pertinent as peer scores/comments were not given to the presenters. Thus, it is a possible fairness issue with scoring, and not how the feedback recipients accept/reject or utilize comments. Even then, the presence of some bias does not invalidate peer assessment as a practice in general; the practical concern is the effect of assessment procedures and stakes.

The conclusions of Li et al. (2022) and Tanaka (2025) are not mutually exclusive. One structured design might produce reliable holistic ratings from several peer raters, and personal relationships might result in small systematic differences in the ratings given in another design. Therefore, the questions of reliability, fairness and

uptake of feedback are related and require separate evidence. Oversight and transparent criteria are especially important where peer scores are used to determine grades.

3.5 Synthesis and specific evidence gap

Selected research indicates that there are varying conclusions regarding teacher and peer feedback based on the criteria used to determine "effective." Oral-performance gains, the quality of comments, the agreement among raters, self-efficacy beliefs, and learner attitudes offer different types of evidence. The principle gap to be filled here is not that teacher–peer research is lacking, but rather that it is hard to integrate claims from different measurement traditions and speaking tasks. Very little of the studies selected measure improvement in speaking, the quality of feedback, and the learners' responses in a manner that allows them to be compared across all three aspects of the study.

Weng et al. (2025) conducted a synthesis of 15 mobile-assisted peer feedback studies conducted between 2014 and 2023 and outlined advantages and drawbacks within the particular technology area. This further reinforces the need to focus on the mode of instruction and demonstrates that the current review research is focused on a limited mode of instruction. The current synthesis maintains a more general teacher–peer point-of-view but does not pretend to cover all studies of feedback. The more careful mapping it offers is what each type of finding can and cannot establish.

IV. DISCUSSION

4.1 Why performance findings differ

Identifying the winning feedback source cannot answer the first research question. The favorable outcomes of teacher feedback are coexisting with the favorable outcomes of peer correction in the EFL studies (Au and Bardakci, 2020; Li and Hu, 2024; Mahvelati, 2021). The more practical understanding is that these interventions do not have the same instructional roles. Educators can provide in-depth language descriptions and relate feedback to course goals. Peers can provide more instances of observing issues, assessing performance, and interacting with evaluation standards. Both possible mechanisms are plausible, and the chosen studies fail to determine which explains every observed gain.

Task and design should also be explained by a causal interpretation. An oral correction study can put emphasis on accuracy and fluency, but a prepared-presentation assessment can incorporate content, organization and delivery. The Thai comparison also did not eliminate any teacher intervention with peer conditions (Wanchid & Charoensuk, 2024). Therefore, discrepancies between reported results might be due in part to what students were told to do, the manner in which feedback was provided, and what was measured, instead of source identity itself. The overall higher-education presentation outcomes provide an insight into the quality of comments and behavior, which cannot be wholesale transferred to the EFL linguistic proficiency.

4.2 Why perceptions should not stand in for learning

The second research question demonstrates a disequilibrium of evidence. The studies of the perception of peer feedback are explored in a number of studies, yet direct data concerning the speaking self-efficacy and the performance tested are focused on Au and Bardakcak (2020). The positive attitude of a learner towards an activity might be a sign of readiness to engage in it; it does not reveal higher fluency or better accuracy. Similarly, unwillingness to be corrected by peers does not mean that learning is not taking place. This distinction is particularly crucial in light of the observed divergence between preferences and speaking outcomes in Mahvelati (2021).

Neither should the lack of a meaningful self-efficacy-performance relationship in a small sample (Au & Bardakçı, 2020) be generalized into a universal null relationship nor be considered a conflict with the perceived value of feedback. Rather, the existing evidence points in favor of independent reporting of beliefs and observed speaking outcomes. Investigations that integrate perception measures that are appropriately validated as well as repeated speaking measures would assist in elucidating whether the two evolve in tandem with various feedback settings.

4.3 Quality, trust, reliability, and bias

The third research question reveals a common conceptual issue: there are cases when the quality of assessments of feedback, rating reliability, and trust are discussed as the same concept. They do not. Feedback-quality criteria are related to the content and form of advice; rater reliability is related to consistency in ratings; learner trust relates to acceptance or trust in the provider. Friendship bias is relevant to systematic distortions of scores. All of them influence the feedback environment, and none of the indicators can replace the others.

This differentiation thwarts two easy conclusions. It would be inappropriate to conclude based on teacher benefits on presentation-feedback rubrics (van Ginkel et al., 2017a) that teachers should bring about higher benefits in all EFL speaking outcomes. It would be just as unreasonable to assume that based on credible peer ratings (Li et al., 2022) recipients will find all peer comments helpful. Structured rubrics, examples of constructive comments, and teacher review are justifiable classroom protections, but their efficacy needs to be verified in the specific course, and not presupposed to have been shown by all studies in this review.

4.4 What the evidence supports

The practical stance that occurs through the synthesis is conditional but not categorical. Those responsibilities of teachers to explain challenging linguistic aspects, and to moderate high-stakes assessment, can be retained, whereas students can take part in formative observations in guided peer activities. This is an educative deduction of the trend of evidence, and not an impact that is experimentally known by integrating all studies. The evidence does not favor having a universal ranking, or that combination of feedback sources is always better than either source alone. The assertions of an integrated approach must then be put forward as a realistic, falsifiable design alternative as opposed to a certainty.

In the case of research, a more precise correspondence between the feedback question and measurement is the key requirement. The comparison between teacher and peer feedbacks must mention that its main result is speaking proficiency, oral-presentation competency, and content of feedback, student beliefs, or fairness of the peer scores. There are a number of outcomes that can be measured by the same study, but the outcomes should not be reduced to one undifferentiated concept of effectiveness.

V. IMPLICATIONS for EFL TEACHING

The review indicates some of the practices EFL teachers should take into consideration, though they have to be modified to fit the local context and measured instead of presumed to be effective. A lesson on speaking can start with a short demonstration on the part of the teacher of what is considered useful feedback. An example is that following a model presentation, learners could apply a small set of criteria to determine one obviously evident strength and one aspect that the speaker could practice before his or her next presentation. The criteria may be intelligibility, coherence, manageable pauses and language accuracy picked on the basis of the instructional aim and not evaluated simultaneously.

Peer comments are more likely to be usable when students have explicit prompts. “You were not fluent” gives little guidance; “the long pause before your second example interrupted the explanation; practice the transition between these two points” identifies the performance feature and a possible next action. The example is a pedagogical illustration, not a quotation from a study. Teachers can review sample comments and correct misunderstandings without taking over every peer exchange. The improvement observed in tutor-guided peer feedback on several quality criteria in the non-EFL study by van Ginkel et al. (2017a) offers a rationale for trialing such preparation, while the EFL assessment study by Li et al. (2022) shows that learners can participate meaningfully in structured evaluation.

The arrangement should be influenced by the assessment stakes. Peer comments can also be employed to inform revision and reflection in formative practice without defining grades. In cases where peer ratings are used as part of marks, moderation, more than a single rater where possible and a shared rubric should be considered by teachers. According to Tanaka (2025), his research reveals the rationale behind considering social relationships in such decisions. Anonymous or self-recorded ratings could be considered; however, the studies reviewed do not prove that any safeguard can completely overcome bias.

Perception of learners also deserves tracking, particularly in the cases when the students are not accustomed to correction by their peers (Mahvelati, 2021). It is possible to teach the teachers to clarify that constructive criticism is about a performance, but not an individual; allow the teacher to challenge comments; and not equate discomfort with failure in teaching. Short self-reflections may document what the learners believed to have been helpful, and a later speaking task may demonstrate whether a particular change occurred or not. These two types of evidence are different and must not be mixed.

Lastly, educators ought to make decisions about feedback on the basis of purpose as opposed to habit. A repeated pronunciation or grammatical issue might require insightful explanation; a prepared presentation might be useful in organized observation of organization and presentation by multiple listeners. It is reasonable to combine these roles, yet the choice must be cross-examined with the levels of student proficiency, classroom size, time, the type of task, and the quality of remarks that students are capable of writing.

VI. LIMITATIONS AND FUTURE RESEARCH

6.1 Limitations

The main restriction is the targeted, but not exhaustive identification of publications. Nine empirical studies were selected based on the relevance of the review; the review does not provide a database-wide record count, independent screening, systematic risk of bias appraisal, nor pooled estimate of effect. Lack of a study in this review does not imply that such a study does not exist. Availability of source records and of full texts was checked; in some cases it was not possible to obtain full text versions of the documents, and therefore, in some cases only published abstracts or parts of the articles were accessible and used for interpretation.

There is limited and mixed-quality evidence. The majority of the nine empirical studies do not focus on general higher-education oral-presentation courses but on EFL courses, while two EFL studies provide contextual evidence of perceptions or peer scoring without making comparisons of instructional outcomes. In the broad sense, only five selected EFL studies provide direct comparisons between teachers and peers and, even in these, the comparison varies from one study to another. The designs included differ in respect of: speaking task, mode of delivery, sample, outcome measures, and whether sources of feedback were completely separated. The studies therefore cannot provide an overall numerical estimation of the effectiveness of teachers versus peers.

The evidence base pertaining to speaking self-efficacy is more limited than that for performance or perceptions. There are also different ways of measuring trust, the credibility of feedback, assessment reliability and quality of comments, which are also measured or inferred from related constructs. The review therefore intentionally does not rank the reliability of individual sources, and does not extrapolate the absence of any significant association in one small sample to the general population. These restrictions weaken and hinder the conclusions.

6.2 Future research

For future research, conclusions may be drawn by comparing the feedback of teachers, trained peer feedback, and untrained peer feedback in the same EFL speaking tasks with the same outcome measures. Separation of feedback conditions and serial evaluations could allow for further sourcing and identify if improvements are maintained after the activity. Studies should report learners' proficiency, and the linguistic or presentation features being measured should be included if possible.

Another priority is to gauge speaking self-efficacy, perceptions of learners' self-efficacy, and performance ratings based on the independent assessment of performance, without reducing those ratings to a confidence score. Feedback credibility needs to be separately investigated in terms of accuracy, specificity of feedback, and trust placed by recipients in the feedback, and also whether the feedback is taken up and the reliability of the ratings. Longitudinal or classroom based designs may compare the attitudes of students before and after they have experience in giving and receiving feedback.

Last, but not least, there is a need for more detailed research on interpersonal relationships and assessment stakes in feedback research. Leniency for friendship matters may differ from the context of peer giving of ungraded suggestions. Further research from secondary-school and underrepresented EFL contexts would also contribute to the question of feasibility and how this is affected by the factors of class size, learner experience and local assessment expectations. The most useful question for future research is not who always provides better feedback, but which feedback procedures will enhance which speaking outcomes for whom and in what manner.

VII. CONCLUSION

The present focused critical narrative review included a selection of literature concerning teacher and peer feedback in EFL speaking, as well as two general studies on the feedback process in higher education classes that focus on oral presentations. The EFL results do not reveal any superior source of feedback. Various aspects of the speaking learning process are captured by measures of perception, speaking self-efficacy, feedback quality, and rating reliability; and reported performance benefits depend on the speaking task and feedback arrangement.

The main take home message is to consider the purpose and context in which feedback is used when making decisions. Explanation by the teacher can deal with language issues that call for expertise and structured peer assessment can provide the learner with an extra agency for noticing, evaluating and reacting to performance. Therefore, the sources are not adequate just on the basis of their name. Focusing on clear criteria, positive comments and ways to use feedback is important, particularly for learners new to peer assessment.

Given that the evidence was obtained in a non-exhaustive search, and is from different settings/designs, the conclusions are qualified. The review does not argue that there is always a better way to use complementary feedback – it does not claim that there is a single correct combination. It is important to make more specific comparisons of speaking improvement, self-efficacy and feedback quality, before more compelling comparisons can be drawn.

REFERENCES

[1]. Au, H. Y. C., & Bardakçı, M. (2020). An analysis of the effect of peer and teacher feedback on EFL learners' oral performances and speaking self-efficacy levels. *International Online Journal of Education and Teaching*, *7*(4), 1453–1468. https://iojet.org/index.php/IOJET/article/view/895

[2]. Charoensuk, V., & Wanchid, R. (2025). Contributions or constraints: Uncovering students' and teachers' perspectives on peer feedback activities in an English oral presentation course. *rEFLections*, *32*(2), 836–856. https://doi.org/10.61508/refl.v32i2.282179

[3]. Li, J., Huang, J., & Cheng, S. (2022). The reliability, effectiveness, and benefits of peer assessment in college EFL speaking classrooms: Student and teacher perspectives. *Studies in Educational Evaluation*, *72*, Article 101120. https://doi.org/10.1016/j.stueduc.2021.101120

[4]. Li, X., & Hu, W. (2024). Peer versus teacher corrections through electronic learning communities and face-to-face classroom interactions and EFL learners' passion for learning, speaking fluency, and accuracy. *Heliyon*, *10*(4), Article e25849. https://doi.org/10.1016/j.heliyon.2024.e25849

[5]. Mahvelati, E. H. (2021). Learners' perceptions and performance under peer versus teacher corrective feedback conditions. *Studies in Educational Evaluation*, *70*, Article 100995. https://doi.org/10.1016/j.stueduc.2021.100995

[6]. Tanaka, M. (2025). Friendship bias in peer assessment of EFL oral presentations. *Assessment in Education: Principles, Policy & Practice*, *32*(5–6), 515–537. https://doi.org/10.1080/0969594X.2025.2570248

[7]. van Ginkel, S., Gulikers, J., Biemans, H., & Mulder, M. (2017a). Fostering oral presentation performance: Does the quality of feedback differ when provided by the teacher, peers or peers guided by tutor? *Assessment & Evaluation in Higher Education*, *42*(6), 953–966. https://doi.org/10.1080/02602938.2016.1212984

[8]. van Ginkel, S., Gulikers, J., Biemans, H., & Mulder, M. (2017b). The impact of the feedback source on developing oral presentation competence. *Studies in Higher Education*, *42*(9), 1671–1685. https://doi.org/10.1080/03075079.2015.1117064

[9]. Wanchid, R., & Charoensuk, V. (2024). Effects of feedback types and English achievement on EFL oral presentation performance of Thai undergraduates. *PASAA*, *69*, 153–194. https://doi.org/10.58837/CHULA.PASAA.69.6

[10]. Weng, J., Noordin, N., Ahmad, N. K., & Chen, L. (2025). A systematic review of mobile-assisted peer feedback in enhancing speaking proficiency for EFL learners. *Cogent Education*, *12*(1), Article 2504228. https://doi.org/10.1080/2331186X.2025.2504228