



The Effect of Task-Based Language Teaching on Speaking Performance and Willingness to Communicate of Indonesian EFL Senior High School Students

Muhammad Ali Imran¹, Nurasia Natsir^{2*}

¹Universitas Muhammadiyah Makassar, Indonesia

²Institut Ilmu Kesehatan Pelamonia Makassar, Indonesia

*Corresponding Author: nurasianatsir87@iikpelamoniamakassar.ac.id

Abstract. Speaking remains one of the most difficult skills for Indonesian learners of English as a foreign language (EFL), partly because classroom instruction is often dominated by teacher talk, grammar explanation, and memorized dialogues that provide few opportunities for meaningful communication. Task-Based Language Teaching (TBLT) has been proposed as an alternative that engages learners in meaning-focused tasks, but empirical evidence from Indonesian secondary schools remains limited. This study investigated the effect of TBLT on students' speaking performance and willingness to communicate (WTC). A quasi-experimental pretest–posttest non-equivalent control group design was employed with 64 eleventh-grade students at a senior high school in [City X], Indonesia. The experimental group ($n = 32$) was taught using Willis's TBLT framework for eight meetings, while the control group ($n = 32$) received conventional Presentation–Practice–Production (PPP) instruction. Data were collected through a speaking test assessed by two raters using an analytic rubric and a WTC questionnaire. The results showed that the experimental group significantly outperformed the control group in the speaking posttest ($M = 74.63$ vs. 65.28 ; $t(62) = 4.57$, $p < .001$, $d = 1.14$), with the largest gains in fluency and content. The experimental group also reported significantly higher WTC after the intervention ($M = 3.56$ vs. 3.02 ; $t(62) = 4.36$, $p < .001$, $d = 1.09$). The findings indicate that TBLT improves not only speaking performance but also learners' readiness to use English. Pedagogical implications for implementing TBLT within the Indonesian curriculum are discussed.

Keywords: EFL; Senior High School; Speaking Performance; Task-Based Learning; Willingness to Communicate.

1. INTRODUCTION

English is a compulsory subject in Indonesian senior high schools and is widely regarded as essential for access to higher education, employment, and international communication (Lamb & Coleman, 2008). Despite years of formal instruction, however, many Indonesian students graduate from secondary school without being able to hold a simple conversation in English. International benchmarks consistently place Indonesia in the low proficiency band among countries where English is not the first language (EF Education First, 2024). Among the four language skills, speaking is frequently reported as the most problematic, because students lack vocabulary, fear making mistakes, and have limited opportunities to use English outside the classroom (Lamb, 2004).

One factor contributing to this problem is the dominance of traditional instructional approaches. In many classrooms, speaking lessons follow a Presentation–Practice–Production (PPP) sequence, in which teachers first present language forms, students then practise them through drills or memorized dialogues, and finally attempt to produce them in controlled activities (Richards & Rodgers, 2014). Although PPP offers a clear structure, critics argue that

it treats language as a set of discrete forms to be accumulated and gives learners few opportunities to engage in genuine, meaning-focused communication (Ellis, 2003; Skehan, 1998). As a result, learners may be able to reproduce correct sentences in controlled practice but struggle to communicate spontaneously.

Task-Based Language Teaching (TBLT) offers an alternative in which the primary unit of instruction is a task an activity in which learners use language to achieve a communicative outcome, with a primary focus on meaning (Ellis, 2003; Nunan, 2004). TBLT is grounded in second language acquisition (SLA) research showing that interaction, negotiation of meaning, and pushed output facilitate language development (Long, 1996; Swain, 1985). A meta-analysis by Bryfonski and McKay (2019) found that TBLT programs had a strong positive effect on learning outcomes compared with non-task-based instruction. However, studies in Asian contexts have also reported challenges in implementing TBLT, including large classes, examination pressure, and teachers' limited understanding of tasks (Butler, 2011; Carless, 2007; Littlewood, 2007).

In addition to performance, recent research has emphasized the importance of affective factors in speaking, particularly willingness to communicate (WTC), defined as "a readiness to enter into discourse at a particular time with a specific person or persons, using a L2" (MacIntyre et al., 1998, p. 547). WTC is considered a key goal of language instruction because learners who are willing to communicate create more opportunities for practice and learning (Peng & Woodrow, 2010; Yashima, 2002). Yet relatively few studies have examined whether TBLT can simultaneously improve speaking performance and WTC among Indonesian secondary school students.

Task-Based Language Teaching

Ellis (2003) identifies four criteria of a task: a primary focus on meaning, some kind of gap (information, reasoning, or opinion), learners' reliance on their own linguistic resources, and a clearly defined communicative outcome. Willis (1996) proposes a TBLT framework consisting of three phases: the pre-task phase, in which the teacher introduces the topic and task; the task cycle, consisting of task performance, planning, and reporting; and the language focus phase, in which learners analyse and practise specific language features that emerged during the task. This framework preserves a focus on meaning while also providing opportunities to attend to form.

Ellis (2009) responds to criticisms of TBLT by clarifying that it does not neglect grammar but rather integrates focus on form within meaning-focused activities. Long (2015) further argues that tasks should be based on learners' real-world communicative needs. In the

context of Indonesian secondary education, which emphasizes student-centred and contextual learning, TBLT appears to align well with curricular goals.

Speaking Performance

Speaking performance is commonly assessed through analytic rubrics that evaluate several dimensions, such as fluency, accuracy, pronunciation, vocabulary, and content or task fulfilment (Brown & Abeywickrama, 2019). In SLA research, oral performance is often described in terms of complexity, accuracy, and fluency (Housen & Kuiken, 2009). Skehan (1998) argues that learners have limited attentional capacity, so tasks and task conditions can influence which dimension of performance is prioritized. For example, planning time before a task tends to enhance fluency and complexity.

Willingness to Communicate

MacIntyre et al. (1998) propose a heuristic pyramid model of L2 WTC, in which situational variables, such as the desire to communicate with a specific person and state communicative self-confidence, rest on more enduring variables, such as motivation, intergroup attitudes, and personality. Yashima (2002) found that international posture and communication confidence predicted WTC among Japanese EFL learners. Peng and Woodrow (2010) showed that classroom environment significantly influenced WTC in Chinese EFL classrooms. These findings suggest that instructional approaches that create a supportive and communicative classroom environment may increase WTC.

Research Questions

Based on the gaps identified above, this study addressed the following research questions: Is there a significant difference in speaking performance between students taught using TBLT and those taught using PPP? Is there a significant difference in willingness to communicate between students taught using TBLT and those taught using PPP? Which components of speaking performance improve most following TBLT instruction?

2. METHOD

Research Design

This study employed a quasi-experimental design with a pretest–posttest non-equivalent control group (Creswell & Creswell, 2018). This design was chosen because random assignment of individual students was not possible in the school context; therefore, two intact classes were assigned to the experimental and control conditions.

Participants

The participants were 64 eleventh-grade students from two classes at a public senior high school in Makassar, Indonesia. One class ($n = 32$; 18 female, 14 male) was assigned as the experimental group and the other ($n = 32$; 17 female, 15 male) as the control group. The students were aged 16–17 years and had studied English for approximately six years. Both classes were taught by the same teacher to minimize teacher effects. Informed consent was obtained from the school and the students' parents.

Instruments

Two instruments were used. The first was a speaking test in which each student performed a two-to-three-minute monologue and a short interaction with the examiner on familiar topics. Performances were audio-recorded and scored independently by two raters using an analytic rubric adapted from Brown and Abeywickrama (2019), covering five components: fluency, accuracy, pronunciation, vocabulary, and content. Each component was scored on a scale of 1–20, giving a maximum total score of 100. Inter-rater reliability, calculated using the intraclass correlation coefficient (ICC), was .87 for the pretest and .89 for the posttest, indicating good agreement.

The second instrument was a WTC questionnaire consisting of 10 items adapted from MacIntyre et al. (1998) and Peng and Woodrow (2010), measuring students' willingness to speak English in various classroom situations (e.g., "I am willing to give my opinion in English during class discussions"). Items were rated on a five-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*). The questionnaire was translated into Indonesian to ensure comprehension. Cronbach's alpha was .84 for the pretest and .86 for the posttest.

Procedure

Both groups took the pretest and completed the WTC questionnaire in the first week. The treatment was then conducted over eight meetings (2×45 minutes each). The experimental group was taught using Willis's (1996) TBLT framework. Each lesson consisted of a pre-task phase (topic introduction, vocabulary activation, and model input), a task cycle (pair or group task performance, planning, and reporting to the class), and a language focus phase (analysis and practice of language features observed during the task). Tasks included information-gap activities, problem-solving tasks, opinion-exchange tasks, and short presentations on topics related to the curriculum, such as giving suggestions, expressing opinions, and describing procedures.

The control group was taught the same topics and language functions using a PPP approach. The teacher presented expressions and grammar, students practised through

repetition and dialogue memorization, and they then produced short dialogues in pairs. After the treatment, both groups took the posttest and completed the WTC questionnaire again.

Data Analysis

Data were analysed using IBM SPSS Statistics. Descriptive statistics (means and standard deviations) were calculated for each group. The normality of the data was checked using the Shapiro–Wilk test, and homogeneity of variance was checked using Levene's test. Independent-samples *t*-tests were used to compare the groups at pretest and posttest, and paired-samples *t*-tests were used to examine within-group changes. Effect sizes were calculated using Cohen's *d*, interpreted as small (0.2), medium (0.5), and large (0.8) (Cohen, 1988). The significance level was set at .05.

3. RESULTS

Speaking Performance

The Shapiro–Wilk test indicated that the speaking scores were normally distributed in both groups ($p > .05$), and Levene's test confirmed homogeneity of variance ($p > .05$). Table 1 presents the descriptive and inferential statistics for the speaking test.

Table 1. Speaking test scores of the experimental and control groups.

Group	Pretest M (SD)	Posttest M (SD)	Gain	Paired <i>t</i> (31)	<i>p</i>
Experimental (TBLT)	58.41 (8.72)	74.63 (7.95)	16.22	15.04	< .001
Control (PPP)	57.94 (9.05)	65.28 (8.41)	7.34	7.16	< .001

At pretest, there was no significant difference between the experimental and control groups ($t(62) = 0.21$, $p = .834$), indicating that the two groups had comparable speaking ability before the treatment. Both groups improved significantly from pretest to posttest. However, at posttest the experimental group scored significantly higher than the control group ($t(62) = 4.57$, $p < .001$), with a large effect size ($d = 1.14$). The gain in the experimental group (16.22 points) was more than twice that of the control group (7.34 points).

Components of Speaking Performance

Table 2 presents the mean scores of each speaking component at pretest and posttest.

Table 2. Mean scores of speaking components (maximum 20 per component).

Component	Exp. pretest	Exp. posttest	Exp. gain	Ctrl. pretest	Ctrl. posttest	Ctrl. gain
Fluency	11.4	15.3	3.9	11.3	12.9	1.6
Accuracy	11.2	13.8	2.6	11.1	12.6	1.5
Pronunciation	11.9	14.6	2.7	11.8	13.2	1.4

Vocabulary	11.8	15.4	3.6	11.8	13.1	1.3
Content	12.1	15.5	3.4	11.9	13.5	1.6
Total	58.4	74.6	16.2	57.9	65.3	7.4

In the experimental group, the largest gains were found in fluency (3.9), vocabulary (3.6), and content (3.4), while the smallest gain was in accuracy (2.6). In the control group, gains were relatively similar across components (1.3–1.6). The difference between groups was therefore most pronounced in fluency, vocabulary, and content.

Willingness to Communicate

Table 3 presents the results of the WTC questionnaire.

Table 3. Willingness to communicate scores (scale 1–5).

Group	Pretest M (SD)	Posttest M (SD)	Gain
Experimental (TBLT)	2.84 (0.52)	3.56 (0.48)	0.72
Control (PPP)	2.86 (0.55)	3.02 (0.51)	0.16

At pretest, there was no significant difference in WTC between the groups ($t(62) = -0.15$, $p = .881$). At posttest, the experimental group reported significantly higher WTC than the control group ($t(62) = 4.36$, $p < .001$, $d = 1.09$). The paired-samples t -test showed a significant increase in WTC in the experimental group ($p < .001$) but not in the control group ($p = .087$).

Item-Level Changes in WTC

Table 4 shows the WTC items with the largest increases in the experimental group.

Table 4. WTC items with the largest increases in the experimental group.

Item	Pretest M	Posttest M	Change
I am willing to speak English in a small group	3.03	3.91	0.88
I am willing to give my opinion in English during discussions	2.66	3.50	0.84
I am willing to ask the teacher questions in English	2.59	3.34	0.75
I am willing to present in front of the class in English	2.41	3.09	0.68

The largest increase was in willingness to speak English in small groups, followed by willingness to give opinions in discussions. Willingness to present in front of the class showed the lowest posttest mean, although it also increased.

4. DISCUSSION

The first research question asked whether there was a significant difference in speaking performance between students taught using TBLT and those taught using PPP. The results showed that the TBLT group significantly outperformed the PPP group, with a large effect size. This finding is consistent with the meta-analysis of Bryfonski and McKay (2019), which

reported strong positive effects of TBLT on learning outcomes. It also supports the interaction hypothesis (Long, 1996) and the output hypothesis (Swain, 1985): during tasks, students negotiated meaning, noticed gaps in their language, and were pushed to produce comprehensible output. In contrast, PPP activities provided practice of predetermined forms but fewer opportunities for spontaneous, meaning-focused interaction.

The third research question concerned the components of speaking performance. The largest gains in the TBLT group were in fluency, vocabulary, and content, while the smallest gain was in accuracy. This pattern is consistent with Skehan's (1998) trade-off hypothesis, which suggests that meaning-focused tasks tend to prioritize fluency over accuracy because learners have limited attentional resources. The planning stage in Willis's (1996) task cycle may have contributed to improved fluency and content, as planning allows learners to organize ideas and retrieve vocabulary. The smaller gain in accuracy suggests that the language focus phase may need to be strengthened, for example through more explicit feedback or post-task activities that draw attention to form (Ellis, 2009).

The second research question asked whether TBLT affected WTC. The results showed a significant increase in WTC in the TBLT group but not in the PPP group. This finding supports the argument that classroom environment influences WTC (Peng & Woodrow, 2010). In TBLT lessons, students worked in pairs and small groups, which reduced the anxiety associated with speaking in front of the whole class and increased their communication confidence, a key antecedent of WTC in the model of MacIntyre et al. (1998). The item-level analysis supports this interpretation: the largest increase was in willingness to speak in small groups, while willingness to present in front of the class remained relatively low. This suggests that WTC develops gradually from lower-risk to higher-risk situations.

These findings contribute to the debate on the feasibility of TBLT in Asian contexts. While previous studies have reported challenges such as large classes and examination pressure (Butler, 2011; Carless, 2007; Littlewood, 2007), this study shows that TBLT can be implemented effectively in an Indonesian senior high school when tasks are aligned with curriculum topics and when the teacher provides adequate pre-task support. The findings are also consistent with Lamb's (2004) observation that Indonesian learners are motivated to use English but need supportive opportunities to do so.

Table 5 compares the present findings with those of previous studies on TBLT and WTC.

Table 5. Comparison with previous studies.

Study	Context	Main finding	Relation to the present study
Bryfonski & McKay (2019)	Meta-analysis, various contexts	Strong positive effect of TBLT on learning outcomes	Consistent: large effect on speaking ($d = 1.14$)
Carless (2007)	Secondary schools, Hong Kong	TBLT feasible but constrained by examinations and discipline	Partly consistent: feasible when tasks align with curriculum
Peng & Woodrow (2010)	University EFL, China	Classroom environment predicts WTC	Consistent: TBLT environment increased WTC
Skehan (1998)	Theoretical / experimental	Trade-off between fluency and accuracy in tasks	Consistent: larger gains in fluency than accuracy
Present study	Senior high school, Indonesia	TBLT improves speaking and WTC	Extends evidence to Indonesian secondary EFL

Compared with these studies, the present study contributes by examining linguistic and affective outcomes simultaneously in an Indonesian secondary school. The large effect size found here is somewhat higher than the average effects reported in meta-analyses, which may reflect the relatively short duration of the study, the strong contrast between TBLT and a highly controlled PPP condition, and the fact that the same teacher taught both groups with high commitment to the intervention. Replication with different teachers is therefore needed to confirm the robustness of the effect.

It is also worth considering why WTC increased alongside performance. One plausible explanation is a reciprocal relationship: as students experienced success in completing tasks, their perceived communicative competence increased, which in turn made them more willing to speak, creating further opportunities for practice. This interpretation is consistent with the situational model of MacIntyre et al. (1998), in which state communicative self-confidence is a direct antecedent of WTC. Future studies using repeated measures could examine this reciprocal relationship more directly.

Pedagogically, the findings suggest that teachers should incorporate meaning-focused tasks into speaking instruction, use pair and group work to build confidence before whole-class performance, provide planning time to enhance fluency and content, and include a systematic language focus phase to address accuracy. Teacher training programs should also help teachers design tasks that fit their students' levels and curricular goals.

This study has several limitations. First, it used intact classes rather than random assignment, so pre-existing differences between groups cannot be completely ruled out, although pretest results showed comparable abilities. Second, the treatment lasted only eight meetings, so long-term effects are unknown. Third, the sample was limited to one school.

Future research should involve larger and more diverse samples, longer interventions, delayed posttests, and qualitative data, such as classroom observations and interviews, to explore how students experience TBLT.

5. CONCLUSION

This study investigated the effect of TBLT on the speaking performance and willingness to communicate of Indonesian EFL senior high school students. The results showed that students taught using TBLT achieved significantly higher speaking scores than those taught using PPP, with the largest gains in fluency, vocabulary, and content. TBLT also significantly increased students' willingness to communicate, particularly in small-group situations. These findings indicate that TBLT is an effective approach for developing both the linguistic and affective dimensions of speaking. Teachers are encouraged to integrate meaning-focused tasks into their teaching while strengthening attention to accuracy through a systematic language focus phase. Future studies should examine the long-term effects of TBLT and explore its implementation in different educational contexts.

REFERENCES

- Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices* (3rd ed.). Pearson.
- Bryfonski, L., & McKay, T. H. (2019). TBLT implementation and evaluation: A meta-analysis. *Language Teaching Research*, 23(5), 603–632. <https://doi.org/10.1177/1362168817744389>
- Butler, Y. G. (2011). The implementation of communicative and task-based language teaching in the Asia-Pacific region. *Annual Review of Applied Linguistics*, 31, 36–57. <https://doi.org/10.1017/S0267190511000122>
- Carless, D. (2007). The suitability of task-based approaches for secondary schools: Perspectives from Hong Kong. *System*, 35(4), 595–608. <https://doi.org/10.1016/j.system.2007.09.003>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- EF Education First. (2024). *EF English Proficiency Index 2024*. <https://www.ef.com/epi>
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19(3), 221–246. <https://doi.org/10.1111/j.1473-4192.2009.00231.x>

- Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461–473. <https://doi.org/10.1093/applin/amp048>
- Lamb, M. (2004). Integrative motivation in a globalizing world. *System*, 32(1), 3–19. <https://doi.org/10.1016/j.system.2003.04.002>
- Lamb, M., & Coleman, H. (2008). Literacy in English and the transformation of self and society in post-Soeharto Indonesia. *International Journal of Bilingual Education and Bilingualism*, 11(2), 189–205. <https://doi.org/10.2167/beb493.0>
- Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching*, 40(3), 243–249. <https://doi.org/10.1017/S0261444807004363>
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press. <https://doi.org/10.1016/B978-012589042-7/50015-3>
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Wiley-Blackwell.
- MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, 82(4), 545–562. <https://doi.org/10.1111/j.1540-4781.1998.tb05543.x>
- Nunan, D. (2004). *Task-based language teaching*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511667336>
- Peng, J.-E., & Woodrow, L. (2010). Willingness to communicate in English: A model in the Chinese EFL classroom context. *Language Learning*, 60(4), 834–876. <https://doi.org/10.1111/j.1467-9922.2010.00576.x>
- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching* (3rd ed.). Cambridge University Press. <https://doi.org/10.1017/9781009024532>
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.
- Willis, J. (1996). *A framework for task-based learning*. Longman.
- Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal*, 86(1), 54–66. <https://doi.org/10.1111/1540-4781.00136>