ENHANCING EFL STUDENTS' SPEAKING CONFIDENCE AND FLUENCY THROUGH ROLE-PLAY CARDS AND PUPPETS: A CLASSROOM ACTION RESEARCH

Dea Apriliana[1]

Muhammad Rafi'i[2]

Faculty of Teacher Training and Education, Universitas Muhammadiyah Sumatera Utara, Medan, Sumatera Utara

apriliadea001122@gmail.com[1]

muhammadrafii@umsu.ac.id[2]

Abstract

This study described how role play cards and puppets were implemented in speaking instruction and examined changes in the speaking confidence and fluency of 20 Grade 6 EFL students at a primary school in Thailand. A quantitative-dominant mixed-methods classroom action research was conducted in a pre-cycle stage and two cycles. Data were collected through classroom observation, a speaking test, and a questionnaire, and analyzed descriptively. Observed confidence rose from 14.65 to 24.10 out of 32, and observed fluency from 13.15 to 22.65. In Cycle II, all students reached the High or Very High category for confidence and 19 of 20 for fluency. Fluency improved more slowly than confidence, and grammar and pronunciation remained the weakest aspects. Questionnaire results were positive but more cautious than the researcher-rated results, so the findings are read as improved observed performance while the media were in use, not as proof of a causal effect. Double scaffolding, with cards giving cognitive support and puppets giving affective support, is proposed as a working explanation, and accuracy appears to need a further, feedback-based layer of support.

Keywords: *Role-play cards, puppets, speaking confidence, speaking fluency, double scaffolding, classroom action research*

Introduction

Speaking is widely regarded as one of the most crucial yet challenging skills in English as a Foreign Language (EFL) learning (Brown, 2001, as cited in Soodmand Afshar & Asakereh, 2016) because it requires learners to process language spontaneously. Two aspects are central to oral communication: confidence, the psychological readiness to speak without excessive fear of judgment (Bandura A, 1997), and fluency, the ability to produce language smoothly without frequent, unnatural pauses (Hughes, 2002). A systematic review of 30 studies across more than fifteen countries, including Indonesia, shows that speaking anxiety and the confidence and fluency deficits it produces are a recurring pattern among EFL learners (Widyawati et al., 2026). Similar difficulties have been reported among undergraduate students in Medan, Indonesia, who were reluctant to speak English, whose speech was strongly shaped by their mother tongue, and who struggled with sounds absent from the Indonesian sound system (Rafii & Rangkuti, 2022).

The same pattern appeared among the 20 Grade 6 EFL students of a primary school in Thailand who took part in this study. In the pre-cycle (baseline) classroom observation, the mean fluency score was 13.15 out of 32 (45% Low, 55% Fair), and the mean confidence score was 14.65 out of 32 (10% Low, 90% Fair); no student reached the High or Very High category. The baseline speaking test agreed, with mean scores of 28.2 for fluency and 36.5 for confidence (0–100) and all students in the Poor category. Students often stayed silent, relied on their first language, and spoke with long pauses, excessive self-correction, and code-switching, while anxiety and fear of making mistakes kept them from participating even when their competence was adequate.

Role play has been shown to improve EFL students' speaking fluency and confidence (Hoa, 2025; Sayow & Marsevani, 2024; Siswanto, 2026) and puppets to reduce speaking anxiety and build self-confidence (Mumtaz Hsb & Ekayati, 2026).

Speaking confidence is defined following (Tridinanti, 2018) as (1) belief in one's own ability, (2) acting independently in making decisions, (3) a positive self-concept, and (4) daring to express opinions, supported by self-efficacy (Bandura A, 1997) and willingness to communicate (Macintyre & Legatto, 2011). Speaking fluency follows (Lopez et al., 2021): (1) speed, (2) pauses, (3) repetitions, and (4) corrections, which correspond to the speed, breakdown, and repair dimensions of oral fluency (Suzuki & Kormos, 2023).

(Siswanto, 2026) found that role play cards used across three action research cycles with vocational high school students in Surakarta, Indonesia, led to steady gains in vocabulary, fluency, and confidence. (Sayow & Marsevani, 2024) reported that two cycles of role play improved the speaking scores and willingness to practice of 20 eighth-grade students in Batam, Indonesia. (Hoa, 2025) found that moving from scripted to unscripted role play produced significant gains among 27 secondary students in Vietnam (Cohen's d = 1.33–2.10), with interviews indicating reduced anxiety. Perception studies are similarly positive, although performance anxiety and limited preparation persist (Ky Nhan, 2024; Widyani et al., 2025). All these studies involved secondary-level or older learners in Indonesia or Vietnam, whereas the present study involves Grade 6 learners in Thailand, who typically have more limited vocabulary and shorter attention spans.

Puppets lower the psychological barriers to speaking. (Rivaldi, 2026) found that hand puppets reduced fear of mistakes among eight anxious eighth-grade students in

Surabaya, Indonesia. (Mumtaz Hsb & Ekayati, 2026) found that finger puppets significantly improved the speaking skills and self-confidence of 30 students in Southern Thailand (mean score from 60.28 to 84.32), attributing the effect to the psychological distance created when students voice the puppet rather than expose themselves. (Wulan et al., 2025) combined puppets with drama among 22 fifth- and sixth-grade students in Jambi, Indonesia, and raised classroom engagement from 71.1% to 92%.

The pairing of the two media in this study rests on scaffolding theory. Scaffolding is temporary, contingent support that enables a learner to carry out a task that would otherwise be out of reach and that is withdrawn as competence grows (Balegh et al., 2025), an idea rooted in Vygotsky's (1978) zone of proximal development. In the speaking studies reviewed above, role play and puppets are treated as separate techniques, and the kind of support each provides is seldom specified. This study distinguishes two kinds. Role play cards give *cognitive and linguistic scaffolding*: they supply a situation, a role, and an expected sequence of turns, so that the learner does not have to invent content and structure while speaking. Puppets give *affective scaffolding*: because the learner speaks through the puppet, personal exposure is reduced, which is the psychological distance described by (Mumtaz Hsb & Ekayati, 2026), and the fear of judgment that undermines confidence is lowered.

Double scaffolding is therefore used in this article as a working conceptual proposition: when cognitive and affective supports are combined, a young learner can attempt to speak with less fear and a lighter planning burden, and the resulting successful attempts are the mastery experiences through which self-efficacy and, in turn, willingness to communicate develop (Bandura A, 1997; Macintyre & Legatto, 2011). Read this way, the proposition implies that confidence and fluency, which depend on emotional safety and on knowing what to say, should respond first, whereas accuracy, which neither support addresses directly, should respond more slowly; and that gains seen while the supports are in place do not by themselves show that learners can speak without them. These implications serve as an interpretive lens in the Discussion; they were not formulated as hypotheses tested in advance, and the study design cannot test the proposition itself.

Against this background, three gaps remain in the studies reviewed. First, role play research has concentrated on secondary-level and older learners, and puppet research has examined puppets alone or alongside drama (Wulan et al., 2025); structured role play cards combined with puppets for upper-primary learners, studied through cyclical classroom action research, have to our knowledge received little attention. Second, existing studies seldom explain why a combination should work, so the mechanisms of reduced exposure and provided structure are mentioned but not specified as complementary supports. Third, most studies report either performance scores or perceptions, and where both are collected, the differences between them are rarely analyzed. This study addresses these gaps by (1) documenting how cards and puppets were used in a Grade 6 EFL classroom in Thailand, (2) interpreting the results through the double-scaffolding proposition, and (3) comparing researcher-rated performance with students' self-reports and treating their discrepancy as data. It answers three research questions: (RQ1) How are role play cards and puppets implemented in speaking instruction? (RQ2) To what extent do students' speaking confidence scores improve when role play cards and puppets are used? (RQ3) To what extent do students' speaking fluency scores improve when role play cards and puppets are used?

RESEARCH METHOD

Research Design

This study used classroom action research (CAR) with a quantitative dominant, concurrent embedded mixed-methods design: numerical data from observation scores, speaking tests, and a questionnaire were the main evidence, supported by qualitative notes recorded during the action. CAR was chosen because the aim was to improve speaking instruction in one classroom through repeated cycles of action and reflection, not to estimate an effect in a population. It followed the spiral model of (Kemmis et al., 2014), in which planning, acting, observing, and reflecting are repeated across cycles. A pre-cycle (baseline) stage of two meetings, held before any use of the cards and puppets, was followed by two cycles, each held over two meetings. As is usual in CAR, there was no control group, so the pre-cycle stage is the only point of comparison. Quantitative data were used to answer RQ2 and RQ3, and qualitative notes to answer RQ1.

Participants

The participants were all 20 Grade 6 EFL students of a primary school in Thailand, selected through total sampling because the aim was to improve speaking instruction in this intact class rather than to generalize to a wider population. Written consent was obtained from the school and the students' parents/guardians prior to data collection. The students were 11 to 12 years old, and their first language was Thai. The researcher delivered the activities, recorded the observations, and administered the speaking tests; the classroom teacher assisted the researcher in assessing the students.

Instruments and Data Collection

Three instruments were aligned with the definitions above (Table 1). First, a classroom observation checklist recorded eight fluency indicators and eight confidence indicators for each student on a 1–4 scale, so that totals ranged from 8 to 32 and were categorized as Low (8–12), Fair (13–19), High (20–26), or Very High (27–32). It also contained seven whole class Yes/No indicators with notes on the effectiveness of the cards and puppets, completed in every meeting. Individual observation scores are reported for the pre-cycle stage and Cycle II only. Second, an individual speaking score sheet combined a five-aspect analytic rubric (fluency, grammar, vocabulary, pronunciation, and discourse/interaction, each scored 1–5) with a six-item confidence checklist (scored 1–4); students were tested in pairs, and scored individually, at the end of the pre-cycle stage and of each cycle, and scores were expressed as a percentage of the maximum and categorized as Excellent (86–100), Good (71–85), Fair (56–70), or Poor (55 or below). Third, a 20-item questionnaire on a three-point scale (Yes = 3, Sometimes = 2, No = 1) had ten confidence items based on (Tridinanti, 2018) and ten fluency items based on (Lopez et al., 2021). It was administered once, at the end of the intervention, and the researcher read each item aloud in English and explained it in Thai. Questionnaire results were not classified into categories; they are reported as mean scores and as the percentage of Yes, Sometimes, and No answers.

Table 1. Summary of the research instruments

Instrument	Content and source	Scoring	When and by whom
Classroom observation checklist	8 confidence indicators (Tridinanti, 2018) and 8 fluency indicators (Lopez et al., 2021) per student; 7 whole-class Yes/No indicators on the cards and puppets	1–4 per indicator; totals 8–32; Low, Fair, High, Very High	Whole class items in every meeting; individual items at pre-cycle and Cycle II; researcher, assisted by the classroom teacher
Speaking score sheet	Analytic rubric (fluency, grammar, vocabulary, pronunciation, discourse/interaction); 6-item confidence checklist	Rubric 1–5 per aspect; checklist 1–4 per item; percentage of maximum; Poor to Excellent	End of pre-cycle, Cycle I, and Cycle II; researcher, assisted by the classroom teacher
Questionnaire	10 confidence items (Tridinanti, 2018); 10 fluency items (Lopez et al., 2021)	Yes = 3, Sometimes = 2, No = 1; totals 10–30 per aspect	Once, at the end of the intervention; read aloud in English and explained in Thai

Validity and Reliability of the Instruments

Content validity. Every item was derived from the operational definitions in the Introduction, with confidence indicators drawn from (Tridinanti, 2018) and fluency indicators from (Lopez et al., 2021), so that each item can be traced to a component of the construct (Table 1). The questionnaire was built from a blueprint that assigns every item to a component of these definitions (belief in one's own ability, acting independently, positive self-concept, and daring to express opinions for confidence; speed, pauses, repetitions, and corrections for fluency) and was worded in simple English for young learners, and the observation checklist and speaking score sheets define each indicator and score category in advance. The instruments were not, however, put through a formal expert-judgment procedure or piloting, so their validity rests on this theoretical alignment and the explicit definition of indicators, and the absence of expert validation is acknowledged as a limitation.

Reliability. The observation checklist and speaking score sheet were applied with fixed scoring descriptors by the researcher, with the classroom teacher assisting in assessing the students. Speaking test performances were given in pairs, in random order, with each student scored individually and live; according to the testing procedure, they were also recorded so that scores could be checked against the recording. No inter-rater agreement was calculated between the researcher and the teacher. The internal consistency of the questionnaire, estimated with (Russ-Eft, 1980) Cronbach's alpha from the 20 students' responses, was $\alpha = .67$ for the ten confidence items and $\alpha = .34$ for the ten fluency items ($\alpha = .74$ for all 20 items). The confidence value is just below the conventional .70 threshold and the fluency value is low, as can be expected for ten items answered on a

three-point scale by only 20 young learners; several items, especially in the fluency part, correlated weakly with the rest of their subscale. The questionnaire results, particularly for fluency, are therefore treated as indicative only. Reading the questionnaire aloud in English and Thai was intended to keep reading ability from affecting the answers. Because the evidence on reliability is limited, the results are interpreted descriptively and with caution.

Data Analysis

Data were analyzed in four steps. First, all instruments were scored as described above, and observation totals and speaking test percentages were assigned to their categories, while questionnaire results were not categorized. All totals and means were recomputed from the item level scores before analysis. Second, quantitative data were analyzed descriptively: mean scores were calculated for each aspect and stage; the gain was calculated as the later mean minus the baseline mean, with the range of individual gains reported; the number and percentage of students in each category were counted; and indicator-level means were examined to see which behaviours changed most and least. For the questionnaire, mean item scores and the percentage of Yes, Sometimes, and No answers were calculated for each aspect and component, and Cronbach's alpha was calculated for each part of the questionnaire. Because the class was small (N = 20), intact, and without a control group, inferential tests were not used, and the differences reported are magnitudes of change in this class, not tests of statistical significance. Third, qualitative notes were analyzed through data reduction, display, and conclusion drawing: notes from each meeting were reduced to statements about the cards, the puppets, and student interaction, grouped by topic, and used to answer RQ1. Fourth, the strands were triangulated by comparing the direction and size of change across the three sources for each aspect and at indicator level, and by recording where they agreed and where they diverged. Because the researcher rated both the observation and the speaking test, only the questionnaire offers a view independent of the researcher, so triangulation is partial.

Research Procedure

The study followed the stages summarized in Table 2. In the pre-cycle stage, before any use of the cards and puppets, students were observed during their regular speaking lesson and took the speaking test over two meetings, and these results identified the problems that the first action addressed. In these meetings, students took part in their regular speaking activities without the cards and puppets while the researcher observed them.

Table 2. Overview of research stages and data collection

Stage	Meetings	Main activities	Data collected
Pre-cycle	2	Regular speaking lesson without the cards and puppets; speaking test	Individual observation; speaking test

Stage	Meetings	Main activities	Data collected
Cycle I	2	Planning; acting (scenario, card and puppet, rehearsal, performance); observing; reflecting on fluency, grammar, pronunciation, and continuing despite mistakes	Whole-class checklist in each meeting; speaking test in meeting 2
Cycle II	2	Slightly more demanding cards and different group composition; same sequence of activities	Whole-class checklist; individual observation; speaking test in meeting 2
End of intervention	–	Questionnaire read aloud in English and explained in Thai	Questionnaire (once)

Planning. In the planning stage of each cycle, the researcher and the classroom teacher prepared the lesson plan, the role play cards (each giving a scenario and a character role), the puppets, and the instruments. Ten cards were prepared for each cycle, and their content differed between Cycle I and Cycle II. Each card presented an everyday scenario and a role, and the cards in Cycle II were slightly higher in difficulty than those in Cycle I. Students worked in pairs or small groups, and the composition of the pairs and groups was changed between cycles so that students practiced with different partners.

Acting. In every meeting, the researcher introduced a scenario and gave each pair or small group a card and a puppet. Students then rehearsed their roles and performed the dialogue by speaking through the puppets; the speaking activity lasted 15 minutes.

Observing. The researcher recorded students' behavior during both meetings of each cycle, using the whole-class checklist and field notes, and gave the speaking test in the second meeting.

Reflecting. After each cycle, the researcher and the teacher identified which aspects of the cards, puppets, or classroom management needed improvement and revised the plan for the next cycle. After Cycle I, the reflection focused on the weak results of the test, namely fluency (five students still Poor), grammar, pronunciation, and continuing despite mistakes. In Cycle II, the cards were slightly higher in difficulty and the groups were composed differently from those in Cycle I. These changes were made after the reflection on Cycle I, to give students a somewhat greater challenge as their confidence grew and to let them practice with different partners.

Indicators of Success

The action was considered successful when (1) mean observed fluency and confidence scores increased from the baseline (13.15 and 14.65 out of 32) and speaking test scores rose in each successive cycle; (2) at least 75% of students reached the High or Very High category in both aspects; and (3) the end-of-intervention questionnaire showed positive self-reported confidence and fluency.

RESULT AND DISCUSSION

Implementation of Role Play Cards and Puppets (RQ1)

In every meeting, the researcher introduced a scenario, gave each pair or group a card and a puppet, allowed time for rehearsal, and asked students to perform the dialogue through their puppets. The whole-class checklist recorded all seven indicators as observed. The cards gave students a clear scenario and role. The puppets helped shy students take part, aroused visible enthusiasm, and made students look more relaxed and less afraid of making mistakes. Practice also moved from teacher–student exchanges toward peer interaction: students practised with friends, stayed mostly on task, and helped one another complete the dialogue. This is consistent with the double-scaffolding proposition, in which the card supplies what to say and the puppet reduces personal exposure, although field notes alone cannot show which support produced which behaviour. The Cycle I test showed where the media alone were not yet sufficient: five students were still Poor in fluency, grammar (2.30 out of 5) and pronunciation (1.95) were weak, and continuing despite mistakes was the lowest-rated confidence behaviour (2.60 out of 4). The reflection focused on these aspects, and all were higher in Cycle II.

Improvement in Students' Speaking Confidence (RQ2)

Observed confidence rose from a mean of 14.65 at the pre-cycle stage to 24.10 out of 32 in Cycle II (+9.45), and every student improved (gains of 6 to 16 points) (Table 3). At baseline, 18 students were Fair and two Low; in Cycle II, 17 (85%) were High and three (15%) were Very High. The largest indicator gains were for using the puppet as a "mask" (1.45 to 2.90), eye contact (1.80 to 3.25), and voice volume (1.95 to 3.30), while the lowest Cycle II indicator was attempting to speak when unsure of vocabulary or grammar (2.65). These figures describe change in researcher-rated behaviour between two time points in one class; the Discussion considers what they can and cannot show.

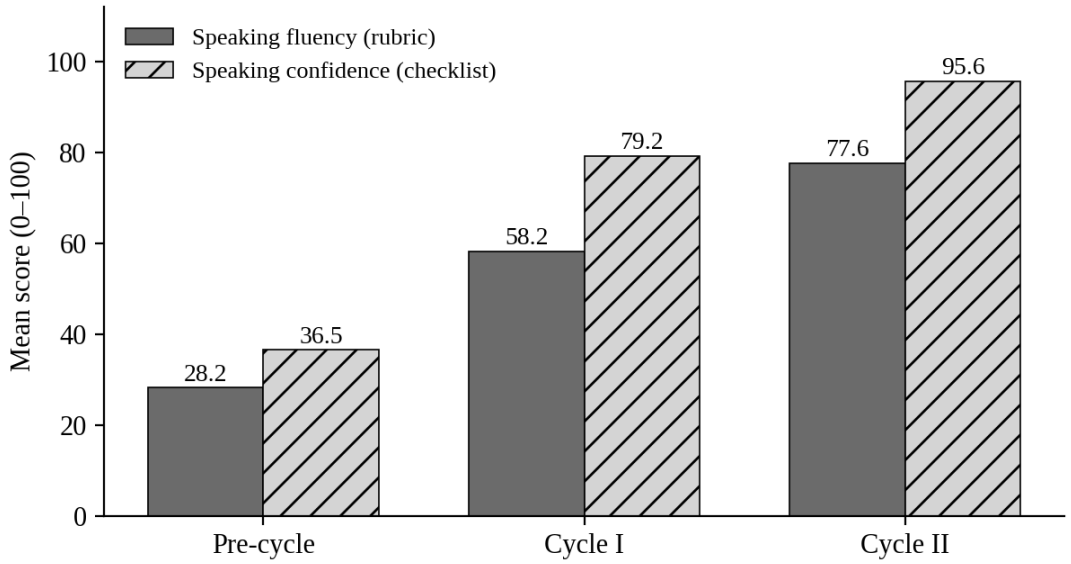
Table 3. Classroom observation results, pre-cycle and Cycle II (N = 20)

Aspect	Stage	Mean (max. 32)	Low	Fair	High	Very High
Speaking confidence	Pre-cycle	14.65	2 (10%)	18 (90%)	0 (0%)	0 (0%)
	Cycle II	24.10	0 (0%)	0 (0%)	17 (85%)	3 (15%)
Speaking fluency	Pre-cycle	13.15	9 (45%)	11 (55%)	0 (0%)	0 (0%)
	Cycle II	22.65	0 (0%)	1 (5%)	19 (95%)	0 (0%)

The speaking test showed the same picture (Table 4, Picture 1). The mean confidence score rose from 36.5, with all students Poor, to 79.2 in Cycle I and 95.6 in Cycle II, when 19 students were Excellent. The six confidence behaviours gained similarly (+2.20 to +2.60 on the 1–4 scale), most for supporting gestures and expressions (1.15 to 3.75) and audible volume (1.35 to 3.90). Continuing despite mistakes stayed the lowest-rated behaviour in both cycles (2.60 and 3.60), so tolerance of mistakes developed most slowly. The Cycle II confidence mean is close to the maximum, so little room for further gain remained and the measure may not separate students at the top of the scale.

Table 4. Speaking test results across cycles (N = 20)

Aspect	Stage	Mean (0–100)	Poor	Fair	Good	Excellent
Speaking confidence	Pre-cycle	36.5	20 (100%)	0 (0%)	0 (0%)	0 (0%)
	Cycle I	79.2	0 (0%)	2 (10%)	15 (75%)	3 (15%)
	Cycle II	95.6	0 (0%)	0 (0%)	1 (5%)	19 (95%)
Speaking fluency	Pre-cycle	28.2	20 (100%)	0 (0%)	0 (0%)	0 (0%)
	Cycle I	58.2	5 (25%)	15 (75%)	0 (0%)	0 (0%)
	Cycle II	77.6	0 (0%)	2 (10%)	17 (85%)	1 (5%)



Picture 1. Mean speaking test scores for fluency and confidence across cycles

Improvement in Students' Speaking Fluency (RQ3)

Observed fluency rose from 13.15 to 22.65 out of 32 (+9.50; Table 3). At baseline, nine students (45%) were Low and 11 (55%) Fair; in Cycle II, 19 (95%) were High, one was Fair, and none remained Low. All students improved (gains of 4 to 14 points). The largest gains were for appropriate pauses (1.45 to 3.10), natural pace (1.85 to 3.15), and spontaneous responses to the partner (1.50 to 2.80); the smallest were for connected speech with linking words (+0.90) and speech that stays understandable despite grammatical errors (+0.95). Smooth self-correction and the latter had the lowest Cycle II means (2.65 each).

The speaking test rose more slowly than for confidence (Table 4): from 28.2, with all students Poor, to 58.2 in Cycle I (15 Fair, five Poor) and 77.6 in Cycle II (17 Good, one Excellent, two Fair). The rubric total rose by 12.35 points out of 25 (individual gains of

7 to 15), from 7.05 to 14.55 and 19.40. The fluency aspect itself had the highest rubric mean (1.45, 3.85, and 4.65), whereas grammar (1.35, 2.30, 3.45) and pronunciation (1.40, 1.95, 3.15) were weakest at every stage; pronunciation hardly changed in Cycle I (+0.55) and gained mainly in Cycle II (+1.20). By Cycle II, 19 students were Excellent in confidence while fluency was mostly Good, so fluency remained the weaker aspect, as at baseline.

Students' Self-Reported Confidence and Fluency

All 20 students completed the questionnaire at the end of the intervention. The mean total was 20.70 for confidence (69.0% of the maximum) and 20.05 for fluency (66.8%). Confidence was rated slightly higher than fluency (Table 5). "Sometimes" was the most frequent answer for both aspects (40.0% and 42.5%). In confidence, the highest-rated statements were choosing one's own words in role play (2.60), feeling happy with the puppet (2.40), and finishing the card task (2.35), and the lowest was wanting to show the role play to friends (1.50; 13 students answered "No"). In fluency, the highest were continuing to talk after a mistake (2.45) and speaking not too slowly (2.30), and the lowest were not stopping too long to think (1.65) and speaking fast with the puppet (1.75). Because the questionnaire was given only once, at the end, it shows how students saw their speaking after the intervention but not whether those perceptions changed.

Table 5. Questionnaire results by aspect (N = 20)

Aspect	Items	Mean (1–3)	Yes	Sometimes	No
Speaking confidence (overall)	**1–10**	**2.07**	**33.5%**	**40.0%**	**26.5%**
– Belief in one's own ability	1–3	2.05	31.7%	41.7%	26.7%
– Acting independently	4–6	1.98	30.0%	38.3%	31.7%
– Positive self-concept	7–8	2.02	37.5%	27.5%	35.0%
– Daring to express opinions	9–10	2.27	37.5%	52.5%	10.0%
Speaking fluency (overall)	**11–20**	**2.00**	**29.0%**	**42.5%**	**28.5%**
– Speed	11–12	2.02	27.5%	47.5%	25.0%
– Pauses	13–15	1.78	15.0%	48.3%	36.7%
– Repetitions	16–17	2.02	37.5%	27.5%	35.0%
– Corrections	18–20	2.20	38.3%	43.3%	18.3%

Achievement of the Indicators of Success

Indicators 1 and 2 were met. Observed fluency and confidence increased from the baseline, speaking test means rose in each successive cycle, and in Cycle II 19 of 20 students (95%) were High in fluency and all 20 (100%) were High or Very High in confidence. Indicator 3 was partially met: the questionnaire was moderately positive but more cautious than the observer-rated results. The action is therefore considered to have met its main targets as defined in this study, with the qualification that the third

indicator was only partly met and that these indicators describe change in one class and not its cause.

Discussion

The three data sources agree on the direction of change: students' speaking confidence and fluency, as rated by the researcher, improved substantially over two cycles of combined role play cards and puppets. This agrees with earlier research on role play(Hoa, 2025; Sayow & Marsevani, 2024; Siswanto, 2026) and extends it to Grade 6 learners in a Thai primary school, although differences in design, learners, and instruments mean the sizes of gains should not be compared directly. The especially large rise in confidence is consistent with the psychological-distancing explanation of puppets (Mumtaz Hsb & Ekayati, 2026), although distancing was not measured directly: using the puppet as a "mask" was among the largest observed gains, and students appeared more relaxed, enthusiastic, and less afraid of mistakes, in line with the engagement reported when puppets were combined with drama (Wulan et al., 2025).

These gains cannot, however, be attributed to the media alone. Because there was no cards-only or puppets-only condition, the study cannot show how much each layer of scaffolding contributed, and because there was no control group, it cannot separate the media from other influences. Several alternative explanations remain plausible: repeated practice and familiarity with testing, since students were tested in pairs, and scored individually, at the end of the pre-cycle stage and of each cycle; the novelty of the puppets, which may have raised enthusiasm temporarily; maturation over the study period; and rater expectancy, since the researcher who delivered the lessons also scored the students, with the classroom teacher assisting, so the observation and the speaking test are not independent sources. In addition, the card content (slightly higher in difficulty in Cycle II) and the group composition differed between Cycle I and Cycle II, so the two cycles were not identical conditions and part of the change between them may reflect these differences, although the more demanding cards make it less likely that the Cycle II gains simply reflect easier material. The near-ceiling confidence results in Cycle II further suggest that the confidence checklist was easy to score highly on. The findings are therefore best read as evidence that students' observed speaking behaviour improved substantially while the two media were in use, in a way that is consistent with the double-scaffolding proposition but does not prove it.

Fluency rose more slowly than confidence, with gains concentrated in speed and continuity rather than accuracy. Seen through the speed, breakdown, and repair dimensions of oral fluency (Suzuki & Kormos, 2023), the data show clear gains in pauses and pace but smaller gains in repair, as smooth self-correction remained among the lowest-rated indicators. A plausible interpretation is that scaffolded content and lower anxiety let students keep talking sooner, whereas accuracy and repair need more explicit attention than the two media alone provided. Continuing after mistakes was likewise the slowest confidence behaviour to develop, echoing the nervousness that perception studies report as a persisting challenge (Ky Nhan, 2024; Widyani et al., 2025).

This pattern refines the double-scaffolding proposition. The cards and puppets supplied content structure and emotional safety, which appear sufficient to keep young learners

talking, but neither gives corrective information about form, and grammar and pronunciation were the weakest aspects at every stage. Scaffolding theory holds that support should respond to what the learner needs at the moment and should be faded as competence grows (Balegh et al., 2025). In these terms, the two media were responsive to learners' affective and structural needs but not to their need for accuracy, and they were never withdrawn. For young EFL learners, double scaffolding may therefore need a third, feedback-oriented layer (brief teacher or peer feedback on grammar and pronunciation) and a plan for fading the card and the puppet so that students eventually speak without them. This is a theoretical suggestion generated by the present data; neither addition was tested, and the study did not check whether the gains persist once the media are removed.

The discrepancy between observation and questionnaire. The questionnaire results were more modest than the observation and speaking test results. All students were High or Very High in observed confidence and 19 of 20 were High in observed fluency, yet the questionnaire means were only 69.0% (confidence) and 66.8% (fluency) of the maximum, and "Sometimes" was the most frequent answer for both aspects. The gap also appears at item level, and in both directions. Pauses were the lowest-rated fluency component in the questionnaire (1.78 out of 3) but among the most improved indicators in the observation (1.45 to 3.10). Observers recorded peer interaction and collaboration, yet wanting to show the role play to friends was the lowest-rated confidence statement (1.50; 13 students answered "No"), and fewer students felt calm with the puppet than enjoyed it. Conversely, continuing to talk after a mistake was among the highest-rated fluency statements in the questionnaire (2.45), although continuing despite mistakes was the lowest-rated observed confidence behaviour.

The data cannot decide which source is more accurate, but four explanations deserve consideration. (a) The sources capture different things: observation records visible behaviour during a supported activity, such as volume, eye contact, and pauses, whereas the questionnaire records how students feel about their speaking in general. Self-efficacy theory suggests that beliefs about one's ability are revised only gradually after repeated success (Bandura A, 1997), so perceptions may lag behind behaviour after two short cycles. (b) The measurement differs: the observation and test were scored by the researcher who taught the lessons, assisted by the classroom teacher, which may have inflated ratings, while young learners may judge themselves strictly against absolute statements; the 4-point and 3-point scales are not directly comparable; "Sometimes" offered an easy middle answer; and the fluency items showed low internal consistency ($\alpha = .34$), which weakens the questionnaire as a measure of fluency. (c) The design limits comparison: the questionnaire was given once, with no baseline, so it cannot show whether self-perceptions improved. (d) Task conditions differ: observed performance took place with cards and puppets, and if questionnaire items refer to speaking without them, lower ratings may reflect a real limit on transfer. The implication is that the evidence is stronger for improved observed behaviour than for improved self-perceived confidence and fluency, which is why the third indicator of success was only partly met. Observed gains and students' own perceptions should therefore be read together, and a claim that students "became confident" should be limited accordingly. Baseline and repeated questionnaires, a second rater, and student interviews would help to explain the gap.

CONCLUSION

This classroom action research examined role play cards and puppets in speaking instruction for 20 Grade 6 EFL students in a primary school in Thailand. Regarding RQ1, the media were implemented through two cycles of planning, acting, observing, and reflecting; the cards gave a clear context, the puppets drew shy students in, and students interacted, stayed on task, and collaborated. Regarding RQ2, observed confidence rose from 14.65 to 24.10 out of 32, all students reached the High or Very High category in Cycle II, and the speaking test mean rose from 36.5 to 79.2 and 95.6. Regarding RQ3, observed fluency rose from 13.15 to 22.65, with 19 of 20 students High, and the speaking test mean rose from 28.2 to 58.2 and 77.6, while grammar and pronunciation remained the weakest aspects. The questionnaire, completed once at the end, was more cautious (confidence 69.0% and fluency 66.8% of the maximum), so the first two indicators of success were met and the third was partially met. Within this one class, the findings suggest that role play cards and puppets together are a promising way to support young EFL learners' observed willingness to keep talking, while accuracy needs further support. Because the study had no control group and relied on researcher-led scoring without an estimate of inter-rater agreement, the findings describe change in this class; they do not show that the media caused it, that each medium contributed equally, or that the gains would carry over to other learners or to speaking without the media. Double scaffolding is offered as a working explanation that still needs to be tested.

LIMITATIONS AND SUGGESTIONS

This study has several limitations. First, it involved one intact class of 20 students without a control group and used descriptive analysis only, so the findings cannot be generalized or attributed causally to the media. Second, it did not separate the contribution of the cards from that of the puppets, so double scaffolding remains a conceptual proposition. Third, the researcher, who also delivered the lessons, scored the observation and the speaking test with the classroom teacher's assistance, no inter-rater agreement was estimated, the internal consistency of the questionnaire was borderline for confidence ($\alpha = .67$) and low for fluency ($\alpha = .34$), the instruments had no formal expert validation, and the two sources were therefore not independent. Card content and group composition also differed between the cycles. Fourth, individual observation scores were reported only for the pre-cycle stage and Cycle II, so the path of change between them is unknown, and the questionnaire was collected only once, at the end, with a three-point scale and no baseline. Fifth, repeated testing and the novelty of the puppets may have inflated the gains, and Cycle II confidence results were close to the maximum. Sixth, there was no follow-up, so it is unknown whether the gains persist without the media. Seventh, because students were tested in pairs, a partner's performance may have influenced an individual's scores. Teachers of young learners can combine role play cards that have clear scenarios with puppets to lower affective barriers, adding brief feedback on grammar, pronunciation, and recovering from errors, introducing whole-class performance gradually, and planning to fade the supports over time. Future studies should compare cards-only, puppets-only, and combined conditions, use independent second raters, administer the questionnaire at baseline and after every cycle, and add interviews and a delayed post-test.

REFERENCES

Balegh, Z., Ahmadi-azad, S., & Yousefi, M. H. (2025). The Effect of Teacher Scaffolding on EFL Learners ' Speaking Ability : Moderated by the Behavioral Inhibition / Activation Systems. *Journal If Studies in Learning and Teaching English*, *14*(1), 1–15.

Bandura A. (1997). Albert Bandura Self-Efficacy: The Exercise of Control. In *W.H Freeman and Company New York* (Vol. 43, Number 9).

Hoa, N. T. (2025). Using Role Play to Enhance Speaking Skills among Secondary School Students: A Case Study at an English Centre. *International Journal of Scientific and Research Publications*, *15*(6), 1–5. https://doi.org/10.29322/ijsrp.15.06.2025.p16202

Hughes, R. (2002). Resensi Bukuteaching and Researching Speaking. *Teaching And Researching (Suwarna)*, *12*(2), 327–340.

Kemmis, S., McTaggart, R., & Nixon, R. (2014). The action research planner: Doing critical participatory action research. In *The Action Research Planner: Doing Critical Participatory Action Research*. https://doi.org/10.1007/978-981-4560-67-2

Ky Nhan, L. (2024). Exploring Students' Perceptions of Role-Playing Activities in Boosting English Speaking Confidence. *International Journal of Innovative Science and Research Technology (IJISRT)*, *9*(10), 825–835. https://doi.org/10.38124/ijisrt/ijisrt24oct708

Lopez, J. I., Becerra, A. P., & Ramirez-Avila, M. R. (2021). EFL Speaking Fluency through Authentic Oral Production. *Journal of Foreign Language Teaching and Learning*, *6*(1), 37–55. https://doi.org/10.18196/ftl.v6i1.10175

Macintyre, P. D., & Legatto, J. J. (2011). A dynamic system approach to willingness to communicate: Developing an idiodynamic method to capture rapidly changing affect. *Applied Linguistics*, *32*(2), 149–171. https://doi.org/10.1093/applin/amq037

Mumtaz Hsb, D. O., & Ekayati, R. (2026). The Influence of Finger Puppet Learning Media to Students' Speaking Skills and Self-Confidence in English Learning. *Journal of English Language and Education*, *11*(1), 471–476. https://doi.org/10.31004/jele.v11i1.1893

Rafii, M., & Rangkuti, R. (2022). Pronunciation Error Done By the Student of Islamic Education in Studying About English Phonetics At the Faculty of Religious Studies in the University of Muhammadiyah Sumatera Utara. *Hikari: Jurnal Bahasa Dan Kebudayaan*, *2*(1), 119–124. https://doi.org/10.37301/hikari.v2i1.17

Rivaldi, L. (2026). INVESTIGATING THE ROLE OF HAND PUPPETS IN REDUCING SPEAKING ANXIETY AMONG JUNIOR HIGH SCHOOL EFL STUDENTS. *Getsempena English Education Journal (GEEJ)*, *13*(1), 32–43.

Russ-Eft. (1980). *Validity and Reliability in Survey Research. Technical Report No. 15.* https://eric.ed.gov/?id=ED279726

Sayow, E. L. M., & Marsevani, M. (2024). Improving high school students' speaking skill through role play strategy: A classroom action research. *English Learning Innovation*, *5*(1), 132–141. https://doi.org/10.22219/englie.v5i1.30366

Siswanto, R. (2026). The Use of Role-Play Cards to Improve Hospitality Vocational School Students' Speaking Skills in Responding Guest Complaint. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, *14*(1), 1294–1299.

Soodmand Afshar, H., & Asakereh, A. (2016). Speaking Skills Problems Encountered by Iranian EFL Freshmen and Seniors from Their Own and Their English Instructors' Perspectives. *Electronic Journal of Foreign Language Teaching*, *13*(1), 112–130. http://e-flt.nus.edu.sg/

Suzuki, S., & Kormos, J. (2023). The multidimensionality of second language oral fluency: Interfacing cognitive fluency and utterance fluency. *Studies in Second Language Acquisition*, *45*(1), 38–64. https://doi.org/10.1017/S0272263121000899

Tridinanti, G. (2018). The Correlation between Speaking Anxiety, Self-Confidence, and Speaking Achievement of Undergraduate EFL Students of Private University in Palembang. *International Journal of Education and Literacy Studies*, *6*(4), 35. https://doi.org/10.7575/aiac.ijels.v.6n.4p.35

Widyani, A. F., Ngafif, A., & Basuki. (2025). Role Play as Pedagogical Tool to Enhance Speaking Skills: Students' and Lecturer's Perception. *Scripta English Department Journal*, *12*(2), 213–222. https://doi.org/10.37729/scripta.v12i2.6414

Widyawati, R., Arif, M., & Hakim, A. (2026). Exploring the Factors of EFL Students' Speaking Anxiety in the English Classroom: Systematic Literature Review. *Journal of English Language Teaching*, *6*(1), 57–71.

Wulan, R., Pratiwi, A. I., & Manar, M. (2025). The Use of Puppet Drama to Engage Speaking Skills of Young Learners. *JELT: Journal of English Language Teaching*, *9*(2), 116. https://doi.org/10.33087/jelt.v9i2.201