

VOCATIONAL-ORIENTED DIGITAL LEARNING TASKS, STUDENT ENGAGEMENT, AND SPEAKING ANXIETY IN EFL CLASSROOMS: EVIDENCE FOR A VOCATIONAL IDENTITY MISMATCH

Merti Anggraini¹, Dian N. Marissa²

¹ Lancang Kuning University; mertianggraini7387@gmail.com

² Lancang Kuning University; dian.marissa@unilak.ac.id

ARTICLE INFO

Keywords:

Keywords 1; Student Engagement
Keywords 2; Speaking Anxiety
Keywords 3; Vocational Identity
Keywords 4; Vocational Education
Keywords 5; Explanatory Sequential Design

Article history:

Received 2026-05-24
Revised 2026-06-12
Accepted 2026-06-26

ABSTRACT

Vocational high school students in EFL classrooms often show low speaking engagement and high anxiety, and tasks rarely combine digital delivery with vocational content. Using a sequential explanatory mixed-methods design, this study tested whether vocational-oriented digital task-based language teaching (VO-DTBLT) improved engagement and reduced anxiety more than conventional instruction. A non-equivalent control-group quasi-experiment compared an experimental group ($n = 30$) receiving digital vocational tasks with a control group ($n = 30$) receiving conventional instruction over two weeks at a public vocational high school in Indragiri Hulu, Riau. Gain scores on student engagement and speaking anxiety questionnaires were analyzed with one-way MANOVA, followed by semi-structured interviews with two purposively selected experimental-group students. Groups were equivalent at baseline. The multivariate effect on the gains was significant, Wilks' $\Lambda = .799$, $F(2, 57) = 7.16$, $p = .002$, but reversed: the control group gained more in engagement and anxiety reduction, with far more variable experimental-group anxiety gains. Interviews linked engagement to whether task content matched students' vocational specialization, consistent with vocational identity theory, whereas control-group students' identity was already secure and self-chosen. Within a short intervention window, VO-DTBLT did not outperform conventional instruction, a pattern better explained by vocational-identity mismatch than by a failure of digital or vocational instruction.

This is an open access article under the [CC BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license.



Corresponding Author:

Merti Anggraini¹

¹ Lancang Kuning University; mertianggraini7387@gmail.com

1. INTRODUCTION

English language teaching in the twenty-first century is increasingly expected to draw on digital technology to build learners' communicative competence and readiness for the

workplace, since technology integration can widen access to authentic, real-world language use and support learner autonomy (Delbio, 2025). In vocational high schools, however, this expectation collides with a persistent mismatch: English instruction is often not organized around the occupational content that vocational students actually need, so classroom tasks read as academic exercises rather than preparation for a trade (Saputro & Hima, 2024). Vocational students tend to hold practical, career-oriented learning goals, and when classroom activities fail to display a clear line to future work, engagement and participation tend to fall (Saputro & Hima, 2024). Two symptoms of this mismatch are especially visible in vocational EFL speaking classes: low behavioral, cognitive, and emotional engagement in speaking activities and pronounced English speaking anxiety rooted in fear of mistakes, low confidence, and fear of negative evaluation (Horwitz et al., 1986). Passive engagement, in particular, has been documented as a distinct classroom pattern rather than a simple absence of effort (Nguyen et al., 2018), and it tends to travel together with anxiety: students who are anxious about speaking participate less, which affords them less practice and feedback, which in turn sustains the anxiety (MacIntyre & Gardner, 1994).

Digital and vocationally contextualized tasks have both been proposed, separately, as ways to break this cycle. Digital learning tasks—video-based projects, simulations, and interactive assignments—have been linked to more constructive and interactive engagement in mixed evidence from higher education (Wekerle et al., 2020), and authentic, real-world tasks have been shown to predict engagement and value beliefs at scale, in a study spanning more than 4,000 secondary students across 129 classrooms (Tan & Nie, 2015). Embedding vocational content into digital tasks—job interviews, workplace simulations, customer-service role-plays—is a natural extension of this logic for vocational learners specifically, on the reasoning that a task which visibly resembles a student's own future job will recruit more genuine motivation than a generic classroom exercise (Hermansah et al., 2025). Recent Indonesian vocational-EFL research has begun testing this logic directly: a digital speaking platform built around local vocational and cultural content improved learners' willingness to engage when the platform's design explicitly managed anxiety alongside practice opportunities (Kurniawan et al., 2026), and blended digital-face-to-face instruction produced speaking gains only when sustained over an extended period (Muhria et al., 2023). AI-augmented digital assessment in vocational EFL classrooms has likewise been shown to affect speaking competency indirectly, through learners' motivation and engagement, rather than as an automatic consequence of the technology itself (Dai et al., 2026).

The instructional-design literature offers a further reason to expect digital and vocational framing to work together rather than separately. Task-based language teaching (TBLT) treats a task as a unit of meaning-focused, outcome-oriented language use, and motivational-design frameworks built on top of TBLT specify the conditions under which such a task is likely to sustain learner effort: attention, relevance, confidence, and satisfaction (the ARCS model). Applying ARCS principles to a digitally mediated, vocationally framed task is not simply a matter of adding a screen to an existing worksheet; it requires that the vocational scenario be genuinely relevant to the specific learner completing it, since relevance is the ARCS component most directly tied to a task's real-world congruence with the learner's own goals. Evidence from flipped, digitally supported EFL writing instruction shows that integrating ARCS-consistent motivational design with digital delivery can improve learner outcomes precisely when the relevance component is well matched to the learner group (Mirzaei et al., 2022). This suggests that a vocational-oriented digital task's success may depend less on the presence of digital and vocational features in the abstract than on how precisely those features are matched to the specific students completing the task—a possibility the present study's qualitative phase was well positioned to test once the quantitative result required an explanation.

What remains comparatively rare is a controlled test of a task that combines both design features—digital delivery and vocational contextualization—against conventional

instruction, using an experimental design that can support a causal claim about engagement and anxiety together, in a single vocational-school sample. Existing research has more often treated digital learning, engagement, and speaking anxiety as separate strands rather than as interacting outcomes of one instructional design. Web-based speaking practice has been shown to reduce anxiety in some contexts but not others, depending on how well the platform's feedback and pacing are matched to learner needs (Bashori et al., 2020), and autonomous motivation has been linked to engagement in general EFL settings through boredom and buoyancy as mediating mechanisms, without reference to vocational content specifically (Wang & Liu, 2022). Neither line of work, nor the vocational-EFL studies reviewed above, has tested engagement and anxiety jointly, as outcomes of the same digitally mediated, vocationally framed task, within a single controlled comparison. This gap motivated the present study: could a vocational-oriented, digitally mediated task-based teaching approach (VO-DTBLT) outperform conventional instruction on both engagement and anxiety simultaneously in an actual vocational classroom rather than a controlled laboratory setting?

At SMKN 1 Pasir Penyu, the vocational high school where this study was conducted, English-speaking instruction has not been systematically redesigned around digital delivery, vocational relevance, or explicit anxiety management (Kurniawati & Lintang Sari, 2024), and classroom observation before the study confirmed the two symptoms described above: a small number of confident students dominate speaking turns while the majority remain reluctant to participate. This context made the school a suitable site for testing whether VO-DTBLT could shift both outcomes simultaneously.

The study was designed, in its original quasi-experimental form, to answer a single confirmatory question: does VO-DTBLT significantly affect students' engagement and speaking anxiety compared with conventional instruction? A one-way multivariate analysis of variance (MANOVA) was planned as the primary hypothesis test, following the logic that engagement and anxiety, though conceptually distinct, are affectively and behaviorally intertwined outcomes of the same speaking task and are therefore best tested jointly rather than in two separate univariate tests. As the results below will show, the MANOVA produced a statistically significant multivariate effect, but in a direction opposite to the study's hypothesis: the control group, not the experimental group, showed the larger gains. A purely quantitative design offers no mechanism for explaining a reversed or null pattern of this kind; it can establish that groups differed, but not why digitally mediated, vocationally framed tasks underperformed a plain conventional lesson within the same short window. For this reason, the study was extended into a second, qualitative phase after the quantitative result was known, in which semi-structured interviews with experimental-group students were used specifically to help explain the unexpected pattern in the gain scores. This two-phase logic—quantitative results first, followed by a targeted qualitative phase designed to explain an anomalous or counter-hypothetical quantitative finding—is the defining feature of an explanatory sequential mixed-methods design (Creswell & Plano Clark, 2018), and it has direct precedent in EFL speaking-anxiety research: Ölmezer-Öztürk and Öztürk (2021) used an explanatory sequential QUAN→qual design in which pre-post anxiety-scale results were interpreted in light of subsequent reflection reports and focus-group interviews. The present study follows the same logic, adapted to a between-groups experimental comparison: the quantitative phase established that something had gone against expectation, and the qualitative phase asked the students most directly affected—those in the experimental group—to explain it in their own words.

Accordingly, this study pursued two connected purposes. First, in the quantitative phase, it examined whether VO-DTBLT produced significantly larger gains in student engagement and significantly larger reductions in speaking anxiety than conventional instruction, using gain scores (posttest minus pretest) as the outcome measures submitted to a one-way MANOVA. Second, in the qualitative phase, it examined how the experimental-group students themselves accounted for their engagement with the vocational digital

tasks, in order to explain the direction and variability of the quantitative pattern. Framing the two phases this way—rather than presenting the interview material as an isolated illustrative anecdote—makes the qualitative evidence do real explanatory work: it is the mechanism through which a counter-hypothetical statistical result becomes pedagogically interpretable.

Theoretically, the study draws on three complementary frameworks. Engagement is treated as a multidimensional construct spanning behavioral, cognitive, and emotional components (Fredricks et al., 2004), each of which can respond differently to a change in task format. Speaking anxiety is understood through Horwitz et al.'s (1986) foreign language classroom anxiety construct and MacIntyre and Gardner's (1994) account of how anxiety narrows the cognitive resources available for language production under performance pressure. Because the qualitative phase pointed toward students' occupational self-concept as the operative explanatory mechanism, the discussion also draws on vocational identity theory (Super, 1980; Meijers, 1998; Skorikov & Vondracek, 2011), which holds that adolescents evaluate the relevance of a learning activity partly through the lens of their emerging occupational identity, so that a task perceived as congruent with that identity draws stronger investment than one perceived as incongruent with it. Practically, the findings are intended to help EFL teachers in vocational settings judge how much adjustment time a digitally mediated, vocationally framed task genuinely needs before it can be expected to outperform familiar conventional instruction, and to flag vocational-content alignment, rather than digital delivery alone, as the design variable most directly implicated in whether such tasks succeed.

This study therefore makes two contributions that extend beyond a single-site quasi-experiment. Methodologically, it demonstrates how an explanatory sequential mixed-methods design can be used after the fact, once a quantitative result runs counter to a study's hypothesis, to convert a null-and-reversed statistical pattern into a pedagogically actionable explanation rather than leaving it as an unexplained anomaly. Substantively, it identifies vocational identity—rather than digital literacy, task difficulty, or general motivation—as the variable most directly implicated in why a vocational digital task succeeds for one student and fails for another within the same classroom, a distinction with direct consequences for how such tasks should be designed for vocationally heterogeneous classes.

2. METHODS

2.1 Research Design

This study used a sequential explanatory mixed-methods design (Creswell & Plano Clark, 2018), consisting of a quantitative phase followed by a qualitative phase that was planned specifically to explain the quantitative result. In the quantitative phase, a non-equivalent control-group quasi-experimental design was used, in which two intact classes were assigned, without individual random assignment, to an experimental condition (VO-DTBLT) and a control condition (conventional instruction). Both groups completed a pretest and a posttest on the same two measures, Student Engagement and Speaking Anxiety. In the qualitative phase, semi-structured interviews were conducted with purposively selected experimental-group students after the quantitative results were known, to help explain a pattern in the gain scores that ran counter to the study's original hypothesis. This QUAN → qual sequence, in which the qualitative strand is subordinate to and dependent on the quantitative result, follows the logic used by Ölmezer-Öztürk and Öztürk (2021) to interpret an EFL speaking-anxiety intervention, and distinguishes the present design from either a purely experimental study or a convergent design in which both strands are collected and analyzed in parallel.

2.2 Setting, Participants, and Sampling

The study was conducted at SMKN 1 Pasir Penyu, a public vocational high school in Indragiri Hulu Regency, Riau, Indonesia. The accessible population was all eleventh-grade (Grade XI) students ($N = 437$, across 14 classes and 8 vocational majors) in the 2026/2027 academic year. Two intact classes were selected through purposive sampling on the basis of comparable English proficiency and classroom conditions: XI ATPH 2 (Agricultural Food Crops and Horticulture, $n = 32$) was assigned as the experimental group, and XI MP (Agricultural Mechanization, $n = 32$) as the control group. After removing incomplete responses, 30 students per group ($N = 60$) were retained for analysis. An independent-samples t -test on pretest scores confirmed that the two groups did not differ significantly on either Student Engagement, $t(58) = -0.40$, $p = .689$, or Speaking Anxiety, $t(58) = -1.80$, $p = .077$, before the intervention began, supporting baseline equivalence despite the absence of individual random assignment.

For the qualitative phase, two students from the experimental group were purposively selected for semi-structured interviews on the basis of their divergent responses to the digital vocational tasks during classroom observation: one enrolled in the Culinary Arts and Food Processing program and one in the Agricultural Machinery program. This small, purposive sub-sample is consistent with the explanatory role the qualitative phase plays in this design—illuminating a mechanism behind the quantitative pattern rather than establishing a representative account of the whole experimental group—and its limits are addressed directly in Section 3.5.

2.3 Instruments

Two questionnaires were administered at pretest and posttest. The Student Engagement (SE) questionnaire comprised 18 items across three dimensions (behavioral, cognitive, and emotional engagement in digital learning) adapted from Fredricks et al. (2004) and Bond and Bedenlier (2019). The Speaking Anxiety (SA) questionnaire comprised 20 items across four dimensions (communication apprehension, fear of negative evaluation, performance anxiety, and anxiety specific to digital and vocational speaking tasks) adapted from Horwitz et al. (1986) and MacIntyre and Gardner (1994). Both instruments used a five-point Likert scale (1 = strongly disagree to 5 = strongly agree), with negatively worded items reverse-scored before analysis. Internal consistency, estimated with McDonald's omega and Cronbach's alpha, was good for the SE scale ($\omega = .854$, $\alpha = .848$) and excellent for the SA scale ($\omega = .953$, $\alpha = .950$), both exceeding the conventional .70 threshold (Nunnally, 1978). Item-rest correlations ranged from .20 to .81 for the SE scale and from .48 to .85 for the SA scale; four SE items showed comparatively low item-rest correlations but were retained because removing them did not meaningfully improve overall reliability. Table 1 summarizes the dimensional structure of both instruments.

Table 1. Instrument Blueprints

Instrument	Dimension	Items	Theoretical basis
Student Engagement (18 items)	Behavioral engagement	1–6	Fredricks et al. (2004); Bond & Bedenlier (2019)
	Cognitive engagement	7–12	
	Emotional engagement	13–18	
Speaking Anxiety (20 items)	Communication apprehension	1–5	Horwitz et al. (1986)

Fear of negative evaluation	6–10	
Performance anxiety	11–15	MacIntyre & Gardner (1994)
Digital-vocational speaking anxiety	16–20	Extended dimension for this study

The interview protocol used in the qualitative phase (Section 2.2) consisted of open-ended, semi-structured questions asking students to describe their experience of the digital vocational tasks, what made them want to participate or hold back, and how the task content related, if at all, to their own vocational program; follow-up probes asked students to elaborate on any comparison they drew, spontaneously or otherwise, between the assigned task content and their own field of study. Each interview lasted approximately 15–20 minutes, was conducted in Indonesian to allow students to express nuance freely, and was translated into English for reporting.

2.4 Procedure

The intervention ran across six meetings over a two-week period: a pretest session, four instructional sessions, and a posttest session. In the experimental group, the four instructional sessions introduced, in sequence, (a) orientation to the vocational task format together with a simple digital speaking task (a video-prompted role-play), (b) a paired digital workplace-conversation task, (c) a project-based task in which students created a short vocational video, and (d) a digital job-related speaking presentation. The control group covered comparable speaking content across the same four sessions using conventional, non-digital methods—textbook-based dialogue practice, drilling, and teacher-led worksheet exercises—without vocational or digital framing. Both groups were taught by the same teacher to hold instructional style constant across conditions as far as possible within an intact-classroom design.

2.5 Data Analysis

Quantitative phase. Gain scores (posttest minus pretest) were computed for each participant on both SE and SA and used, rather than raw posttest scores, as the dependent variables submitted to hypothesis testing. This choice follows methodological guidance that, in pretest–posttest designs where groups are not perfectly equivalent at baseline, analyzing change scores removes each participant's own starting level from the outcome and provides a more direct answer to how much a group changed than analyzing posttest scores alone (Vickers & Altman, 2001), and that gain scores remain an interpretable index of change provided the underlying instrument is reliable (Dimitrov & Rumrill, 2003), a condition met by both scales here (Section 2.3). All analyses were conducted in JASP (JASP, 2026).

Before hypothesis testing, the assumptions of one-way MANOVA were checked. Multivariate normality of the two gain scores was supported by a non-significant Shapiro–Wilk test, $W = .973$, $p = .153$. Multicollinearity, assessed with a Pearson correlation between the two gain scores, was not a concern: $r = -.006$, $p = .962$, indicating the two dependent variables were essentially independent and appropriately combined in a single multivariate test rather than being redundant measures of one construct. Homogeneity of the between-group covariance matrices, tested with Box's M , was violated, $\chi^2(3) = 15.13$, $p = 0.002$; because Wilks' Lambda is comparatively robust to this violation when group sizes are equal, as they were here ($n = 30$ per group), the MANOVA proceeded on Wilks' Lambda, with the violation reported transparently as a

limitation (Section 3.5) rather than treated as immaterial. The one-way MANOVA tested the effect of instructional condition (VO-DTBLT vs. conventional) on the combined GAIN_SE and GAIN_SA scores. Follow-up univariate comparisons for each gain score separately were computed as independent-samples t-tests derived from the reported group means and standard deviations (with F equivalent to t^2 for a two-group comparison), using Welch's correction given the unequal group variances flagged by the Box's M result; this derivation is reported explicitly here because a dedicated univariate ANOVA table was not separately generated in JASP.

Qualitative phase. Interview responses were transcribed and analyzed using a content-analysis approach: responses were read for recurring ideas linking task content to willingness to engage, and these ideas were then interpreted against the vocational identity framework (Super, 1980; Meijers, 1998; Skorikov & Vondracek, 2011) that had already been identified, from the quantitative pattern and classroom observation, as the most plausible explanatory candidate. Integration of the two phases followed a data-transformation logic consistent with explanatory sequential design: the quantitative gain-score pattern defined what needed explaining, and the qualitative theme supplied the explanation, which is reported and discussed jointly in Section 3.4.

2.6 Ethical Considerations

Institutional permission to conduct the study was obtained from the school administration of SMKN 1 Pasir Penyu prior to data collection. Students in both groups were informed, through their homeroom and English teachers, that the lessons were part of a research study, that participation in the questionnaires and, for the two interviewees, the follow-up interview was voluntary, and that individual responses would be reported anonymously using participant numbers rather than names. No student was withheld instructional content as a result of group assignment; both groups covered the same speaking topics and received the same number of instructional sessions, differing only in whether the sessions used vocational-oriented digital tasks or conventional methods.

3. FINDINGS AND DISCUSSION

3.1 Quantitative Phase: Descriptive Statistics

Table 2 and Table 3 summarize the pretest, posttest, and gain scores for both groups on Student Engagement (SE) and Speaking Anxiety (SA). The two groups began from a similar baseline on SE (control M = 3.24, SD = 0.28; experimental M = 3.27, SD = 0.30) and on SA (control M = 3.34, SD = 0.52; experimental M = 3.57, SD = 0.49), consistent with the baseline-equivalence test reported in Section 2.2. By posttest, the control group's mean SE score had risen to 3.75 (SD = 0.40), a gain of 0.51, while the experimental group's mean SE score rose only to 3.34 (SD = 0.41), a gain of 0.07. On SA, where lower scores indicate lower anxiety, the control group's mean fell from 3.34 to 3.07, a reduction of 0.27, while the experimental group's mean fell from 3.57 to 3.43, a smaller reduction of 0.14. The experimental group's SA gain scores were also considerably more dispersed (SD = 0.78, range -1.95 to 1.30) than the control group's (SD = 0.37, range -1.05 to 0.60), indicating that the digital vocational tasks moved individual students in different directions rather than producing a uniform reduction in anxiety.

Table 2. Student Engagement (SE) Scores by Group

Group	1	Pretest M (SD)	Posttest M (SD)	Gain M (SD)
-------	---	-------------------	--------------------	----------------

Control	3	3.24 (0.28)	3.75 (0.40)	0.51 (0.44)
Experimental	3	3.27 (0.30)	3.34 (0.41)	0.07 (0.49)

Table 3. Speaking Anxiety (SA) Scores by Group

Group	n	Pretest M (SD)	Posttest M (SD)	Gain M (SD)
Control	3	3.34	3.07	-0.27
Experimental	3	3.57	3.43	-0.14

3.2 Quantitative Phase: Hypothesis Testing

The one-way MANOVA testing the effect of instructional condition on the combined gain scores (GAIN_SE, GAIN_SA) was statistically significant, with Wilks' $\Lambda = .799$, $F(2, 57) = 7.16$, $p = .002$, and a multivariate effect size ($1 - \Lambda$) of .201, a large effect by conventional benchmarks. Given the direction of the group means reported above, this significant multivariate effect reflects a pattern in which the control group outperformed the experimental group on both outcomes, the reverse of the study's original hypothesis.

Follow-up comparisons for each gain score, computed from the reported group descriptives (Section 2.5), indicated that the two groups differed significantly on GAIN_SE, $t(58) = 3.70$, $p < .001$ (partial $\eta^2 = .19$), a large effect, but not significantly on GAIN_SA, $t(58) = -0.80$, $p = .425$ (partial $\eta^2 = .01$), a small effect. This pattern indicates that the significant multivariate result was driven primarily by the engagement gain, where the control group's advantage was consistent and large, and much less by the anxiety gain, where the group means were closer together but the experimental group's much greater variability (Table 3) meant that any average difference between groups was diluted by inconsistent individual responses within the experimental condition itself. Table 4 summarizes the multivariate result, and Table 5 the follow-up comparisons.

Table 4. One-Way MANOVA: Multivariate Test Results

Effect	Wilks' Λ	F	df	p
Intercept	0.64 6	1 5.61	2, 57	< .001
Condition (VO-DTBLT vs. Conventional)	0.79 9	7 .16	2, 57	.002

Table 5. Follow-Up Comparisons by Outcome (Welch-corrected)

Outcome	t	df	p	Partial η^2
GAIN_SE	3.70	5 7.4	< .001	.19

GAIN_SA	-	4	.42	
	0.80	1.1	5	.01

Taken on its own, this quantitative result answers the study's original confirmatory question in the negative: within this two-week window, VO-DTBLT did not produce larger engagement gains or larger anxiety reductions than conventional instruction; if anything, the reverse occurred for engagement, and the experimental group's anxiety response was too variable to reach significance. What the MANOVA cannot say is why a task explicitly designed to be more relevant and more interactive than conventional instruction underperformed it, or why individual experimental-group students responded so differently from one another on speaking anxiety. This is the gap the qualitative phase was designed to address.

3.3 Qualitative Phase: Explaining the Engagement Gap

The wide variability in the experimental group's anxiety gain, together with its smaller engagement gain, suggested that the digital vocational tasks did not land the same way for every student in that condition. To understand why, semi-structured interviews were conducted with two experimental-group students whose classroom behavior during the intervention had been visibly different from one another: one enthusiastic during the tasks and one comparatively withdrawn. Both students had comparable prior exposure to smartphones and digital tools in daily life, and neither described difficulty operating the digital task platform itself; what distinguished them was not digital readiness but the relationship between the task content assigned to their class and their own field of study.

This is a meaningful contrast with the technical- and cognitive-load explanation that might otherwise be assumed for any digital-task underperformance: if digital unfamiliarity alone were driving the engagement gap, both students, who reported similar comfort with digital tools, should have responded similarly. Their vocational backgrounds turned out to be the salient difference between them, not their digital confidence or general classroom demeanor. Both students, without being asked directly about their vocational program, spontaneously framed their willingness to participate in terms of whether the task content matched their own field of study.

Student 10 (Culinary Arts and Food Processing program): “If this interview were about running a bread business, I think I would be more enthusiastic, ma'am.”

Student 15 (Agricultural Machinery program): “If the interview were held at a workshop specializing in agricultural machinery, I think I would be more active in the interview.”

Both statements locate the source of engagement or disengagement in the same place: not in the digital format of the task, nor in its difficulty, but in whether its vocational framing matched the student's own occupational specialization. Student 10, training in food processing, imagined herself more engaged if the task content had been about a bread business; Student 15, training in agricultural machinery, imagined himself more engaged if the setting had been a machinery workshop. Neither student described the technology itself as the obstacle. This convergence across two students from different vocational tracks, expressed in almost identical conditional language (“if this were about my field, I would be more active”), points toward vocational identity—the stability and clarity with which a person identifies with a given occupational role (Super, 1980)—as the operative variable behind the engagement gap, rather than digital literacy or general motivation.

This pattern is consistent with the way vocational identity is theorized to shape how students process a learning task. Meijers (1998) argues that vocational identity develops through an ongoing process of narrative meaning-making in which learners interpret new experiences, including classroom tasks, in relation to an emerging occupational self-concept: a task that resonates with that self-concept is processed as personally relevant, while a task that falls outside it is processed as an external demand carrying little personal stake. Skorikov and Vondracek (2011) add that during adolescence, occupational identity functions as an interpretive lens through which students judge the relevance of an activity, so that a task perceived as vocationally congruent recruits stronger cognitive and emotional investment than one perceived as incongruent. The two responses above illustrate this mechanism directly: the vocational-oriented digital tasks used in the study were fixed and uniform across the whole experimental class, but each student's specialization differed, so a single set of vocational scenarios could be congruent with some students' occupational identity and incongruent with others' within the very same lesson.

A comparison with the control group helps explain why this mismatch mattered less outside the experimental condition. Control-group students, who received no vocational digital framing at all, nonetheless posted a considerably larger engagement gain than the experimental group ($M = 0.51$ vs. $M = 0.07$). A plausible reading, consistent with the vocational identity framework above, is that control-group students in the Agricultural Mechanization program had already chosen that program as their first-choice track, giving them a comparatively stable and secure vocational identity that did not depend on whether any single classroom activity happened to reference their trade. Because their sense of occupational self was not contingent on the vocational framing of a particular task, these students could sustain engagement across a range of instructional formats, including a conventional one with no vocational content at all. This suggests that instructional method and vocational-task alignment matter most precisely for students whose vocational identity is comparatively less secure or less closely matched by the specific tasks in front of them, while students with an already firm, self-selected vocational identity are less sensitive to how any one lesson is framed.

Both groups converged on one further point that helps explain engagement independently of the vocational-mismatch account above. In interviews, students in both conditions reported that they preferred learning English through media-assisted activities because it let them resolve vocabulary gaps quickly, using a phone to look up or construct a sentence rather than pausing to consult a printed dictionary. This preference is consistent with research on mobile-assisted vocabulary learning showing that phone-based lookup tools lower the effort cost of resolving a lexical gap mid-task, sustaining task momentum in a way that a paper dictionary search tends to interrupt (Stockwell, 2010), and with observations that mobile reference tools give learners immediate, in-context vocabulary support without requiring them to break away from a communicative task (Godwin-Jones, 2011). Because this preference appeared in both conditions regardless of group assignment, it functions as a shared facilitating factor rather than an explanation for the between-group difference in gain scores; the vocational-identity mismatch discussed above remains the more direct explanation for why the two groups diverged. Together, these two qualitative findings—a between-group difference explained by vocational-identity alignment and a shared, group-independent preference for media-assisted vocabulary support—illustrate why the qualitative phase was necessary in this design: the first explains the direction of the quantitative result, while the second guards against over-attributing every difference in the data to the treatment itself, since some facilitating conditions clearly operated regardless of which instructional condition a student was assigned to.

3.4 Discussion

Read together, the two phases converge on an account of the study's findings that neither phase could produce on its own. The quantitative phase established that VO-DTBLT, within a two-week window, did not outperform conventional instruction on engagement, and produced an anxiety response too variable across individual experimental-group students to register as a significant group-level reduction. The qualitative phase located the source of that variability in a task-readiness mismatch operating through vocational identity: engagement rose or fell with whether a given task's vocational framing matched a given student's own specialization, a factor that varies within a class of mixed vocational majors even when the digital format is held constant.

This account is compatible with, and extends, several strands of prior work. The methodological logic of pairing a quantitative anomaly with an explanatory qualitative phase also has direct precedent: Ölmezer-Öztürk and Öztürk (2021) used the same explanatory sequential structure to interpret a foreign-language speaking-anxiety intervention, finding through focus-group interviews that students' own accounts of increased comfort and confidence, not visible in the pre-post anxiety-scale means alone, explained why the intervention worked for some participants and not others. The present study reaches a structurally similar conclusion from the opposite direction: rather than confirming a hypothesized improvement, the qualitative phase here explains why an expected improvement did not appear, but in both cases the qualitative material supplies a mechanism that the quantitative scale scores could not reveal on their own, based on students' own account of what the task meant to them.

Kurniawan et al. (2026) found that a digital speaking platform's benefit for vocational learners depended heavily on how well the platform's design managed anxiety alongside practice opportunities, rather than on digital delivery alone—consistent with the present finding that digital format was not, by itself, what students identified as helping or hindering their engagement. Muhria et al. (2023) reported speaking-performance gains from blended instruction only when digital and face-to-face components were combined systematically over an extended period, suggesting intervention length as a moderator that the present study's two-week window may not have allowed for. Dai et al. (2026) similarly found that in AI-augmented vocational EFL classrooms, motivation and engagement mediate rather than automatically follow from digital instruction, which matches the uneven, individually variable engagement and anxiety pattern documented here. None of these studies, however, tested vocational-content alignment directly as an explanatory variable; the present qualitative phase supplies that missing link by showing, in students' own words, that alignment between task content and vocational specialization—not digital novelty, and not general motivation—was the dimension students themselves invoked to explain their own engagement.

The finding also has a bearing on how a task-based digital intervention should be built for a vocationally heterogeneous class. Lambert et al. (2017) note that task engagement tends to be stronger when learners have some ownership over task content; the fixed, externally supplied vocational scenarios used in the present intervention offered no such ownership, and a single set of scenarios could not be simultaneously congruent with the occupational identities of students training in food processing, agricultural machinery, and other majors represented in the same classroom. This points to a concrete design implication distinct from simply extending the intervention period: a vocational digital task designed for a mixed-major vocational class may need to let students select or adapt the occupational scenario to their own specialization, rather than assigning one scenario to an entire class, if it is to recruit the same kind of identity-congruent engagement that a self-chosen vocational program appears to provide on its own, as it did for the control group in this study.

Self-determination theory offers a complementary reading of the same pattern. Deci and Ryan (2000) argue that sustained motivation depends on learning conditions that support autonomy, competence, and relatedness. A vocational task that matches a student's own occupational identity plausibly supports competence and relatedness simultaneously—competence, because the student can draw on prior vocational knowledge to complete the task, and relatedness, because the task situates English use inside a social role (baker, technician) the student already values. A mismatched vocational task, by contrast, offers neither advantage nor adds an unfamiliar digital procedure on top of unfamiliar or irrelevant content, which is consistent with why some experimental-group students appeared to experience the same digital novelty as pressure rather than motivation. This reading treats vocational identity not as a rival explanation to self-determination theory but as the mechanism through which a vocational task does or does not satisfy competence and relatedness for a given learner, which is precisely the kind of learner-by-task interaction that a purely quantitative, group-level comparison cannot detect.

From a student engagement standpoint, these findings extend Fredricks et al.'s (2004) multidimensional account by suggesting that, in vocational settings specifically, the behavioral and emotional components of engagement with a task may be conditioned less by the task's format (digital versus conventional) than by its fit with a moderating individual-difference variable (vocational identity) that is largely external to the task design itself. For speaking-anxiety theory, the wide dispersion in the experimental group's anxiety gain scores is consistent with MacIntyre and Gardner's (1994) account of anxiety as a resource-competing process: for students whose task content matched their vocational identity, the digital format may have freed cognitive resources for language production by heightening relevance, while for students facing an incongruent task, the same digital novelty appears to have added a further source of pressure on top of ordinary speaking anxiety, a pattern also broadly consistent with prior evidence that digitally mediated speaking tools help anxiety only when their relevance and support features are well matched to the learner (Bashori et al., 2020; Mirzaei et al., 2022).

Pedagogically, these findings suggest two distinct and non-substitutable adjustments for vocational EFL teachers considering digitally mediated, vocationally framed speaking tasks. First, a two-week intervention window is very likely too short for any digital task, vocationally framed or not, to overcome the procedural cost of learning a new task format on top of new language content; teachers should expect, and plan around, an adjustment period before comparing outcomes against well-practiced conventional instruction. Second, and more specific to the present findings, vocational alignment cannot be assumed to hold uniformly across a vocationally mixed classroom: a single set of vocational scenarios, however well designed, will be congruent for some students' occupational identity and incongruent for others', and task design should build in some mechanism — student choice of scenario, small-group task variants matched to major, or scenario banks indexed by program — for narrowing that mismatch rather than treating “vocational” as a single, uniform property of a task.

At the institutional level, these findings caution against adopting digital, vocationally themed EFL materials on the assumption that “vocational” is a single property a task either has or lacks. A vocational high school typically enrolls students across several distinct majors within one English class, as was the case in both the experimental and control groups here, and a materials-adoption decision made at the school or curriculum level will only reach its intended effect if it also addresses how a single set of vocational tasks can remain relevant across that internal diversity. This is a design and procurement question as much as a pedagogical one: school policymakers considering similar materials should ask not only whether the materials are vocationally themed, but whether they include enough scenario variety, or enough flexibility for

teachers and students to adapt scenarios, to serve a vocationally mixed classroom rather than a single-major cohort.

Finally, the reversed direction of the quantitative result is itself worth dwelling on, because it is easy to misread a control group's advantage as evidence that conventional instruction is simply superior. The more defensible reading, given the qualitative evidence, is narrower: within this two-week window, a familiar instructional format carried no additional procedural cost, so control-group students could direct their full attention to the speaking content itself, whereas experimental-group students had to divide attention between an unfamiliar digital procedure and content whose vocational relevance varied from student to student. Once the procedural cost of the digital format is amortized over a longer implementation, and once the vocational content is allowed to vary with the learner rather than being fixed for the whole class, there is no theoretical reason to expect the same reversal to persist; the present findings identify the conditions under which VO-DTBLT underperforms, not a ceiling on what vocational digital tasks can achieve.

3.5 Limitations

Four limitations qualify these conclusions. First, group assignment followed intact classes rather than individual random assignment, a common constraint in classroom-based educational research that limits the causal strength of the between-group comparison despite the confirmed baseline equivalence on both outcome measures. Second, the intervention spanned only six meetings across two weeks, which may not have allowed the experimental group to move past an initial technology-adjustment phase; a longer implementation window is needed to determine whether the reversed pattern reflects a genuine limitation of vocational-oriented digital tasks or an artifact of this study's short timeframe. Third, Box's M indicated a violation of the homogeneity-of-covariance-matrices assumption; although Wilks' Lambda is comparatively robust to this violation under equal group sizes, the result should be interpreted with this caveat in mind, and the follow-up univariate comparisons reported in Section 3.2 were derived from group-level descriptive statistics rather than from a dedicated univariate ANOVA output, a transparency the reader should weigh when evaluating the precision of those figures. Fourth, and most directly relevant to the qualitative phase, only two of the thirty experimental-group students were interviewed; while their responses converged clearly on a single, theoretically coherent explanation and were corroborated by the pattern in the quantitative data, they cannot establish how widely the task-readiness mismatch generalizes across the whole experimental group, and future replications of this design should interview a larger and more systematically sampled subset of participants, ideally stratified by vocational major, before treating the explanation as confirmed at the level of the full sample.

4. CONCLUSION

This study set out to test whether vocational-oriented digital task-based language teaching (VO-DTBLT) could outperform conventional instruction in raising engagement and lowering speaking anxiety among vocational EFL learners, and, because the quantitative result ran counter to that expectation, used a second qualitative phase to explain why. The one-way MANOVA showed a significant multivariate effect of instructional condition on the combined gain scores, Wilks' $\Lambda = .799$, $F(2, 57) = 7.16$, $p = .002$, but the direction was reversed from the study's hypothesis: the control group, receiving conventional instruction, gained more in engagement and reduced speaking anxiety by more than the experimental group, whose anxiety response was also markedly more variable across individuals. Within the two-week window and intact-classroom conditions of this study, VO-DTBLT did not outperform conventional instruction on either outcome.

The qualitative phase clarified what a quantitative comparison alone could not: interviews with two experimental-group students converged on task-readiness mismatch, operating through vocational identity, as the mechanism behind both the smaller engagement gain and the wide dispersion in anxiety response. Students engaged more readily with digital vocational tasks when the occupational scenario matched their own specialization and less readily when it did not, while control-group students, whose vocational identity was already anchored in a self-chosen program, sustained engagement independently of any single task's framing. This explanation reframes the study's null-and-reversed quantitative pattern: the issue was not that digital or vocational instruction is ineffective for this population, but that a single, uniform vocational scenario cannot be simultaneously congruent with the occupational identities of students drawn from different majors within one classroom, and that a two-week window may be too brief for any new task format to clear the procedural cost of adjustment.

For classroom practice, these findings suggest that teachers introducing digitally mediated, vocationally framed speaking tasks should plan for a longer adjustment period than a single short unit allows, scaffold students' digital literacy alongside the language content rather than assuming digital fluency, and, where a class draws students from more than one vocational major, build in some mechanism for varying or student-selecting the vocational scenario rather than assigning one scenario uniformly. For theory, the results extend engagement and speaking-anxiety frameworks by identifying vocational identity as a moderating variable that can determine whether a task format that looks promising on paper—combining authenticity, digital interactivity, and vocational relevance—actually translates into stronger engagement and lower anxiety for a given student.

Future research should extend the intervention period to determine whether the reversed pattern reported here persists once students have had more time to adjust to the digital task format, and should interview a larger, systematically sampled subset of experimental-group students, ideally stratified by vocational major, to test how far the task-readiness-mismatch explanation generalizes beyond the two illustrative cases reported here. A design that allows students some ownership over which vocational scenario they complete, rather than assigning a single fixed scenario to a vocationally mixed class, would also make it possible to test directly whether narrowing the vocational-identity mismatch, rather than simply extending time on task, is what closes the gap with conventional instruction.

More broadly, this study illustrates the value of treating an unexpected or counter-hypothetical quantitative result as an invitation to explain rather than an outcome to be reported and set aside. Had the quantitative phase stood alone, the honest conclusion would have been that VO-DTBLT failed to outperform conventional instruction, with little basis for saying why or what to change. The addition of a targeted, theoretically guided qualitative phase converted that null-and-reversed result into a specific, testable account—task-readiness mismatch operating through vocational identity—that gives both researchers and classroom teachers a concrete variable to manipulate in future iterations of vocational digital task design.

ACKNOWLEDGMENTS

: The authors thank the principal, teachers, and students of SMKN 1 Pasir Penyu for their permission and participation in this research.

CONFLICTS OF INTEREST

: The authors declare no conflict of interest.

REFERENCES

Bandura, A. (1997). Self-efficacy: The exercise of control. W.H. Freeman.

- Bashori, M., van Hout, R., Strik, H., & Cucchiarini, C. (2020). Web-based language learning and speaking anxiety. *Computer Assisted Language Learning*, 35(5–6), 1058–1089. <https://doi.org/10.1080/09588221.2020.1770293>
- Bond, M., & Bedenlier, S. (2019). Facilitating student engagement through educational technology: Towards a conceptual framework. *Journal of Interactive Media in Education*, 2019(1), 1–14. <https://doi.org/10.5334/jime.528>
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.
- Dai, X., Wang, Y., Yu, J., Qiu, B., Wang, R., & Na, M. (2026). Motivation and engagement as pathways: How AI-augmented online assessment shapes English-speaking competency in vocational EFL classrooms. *Frontiers in Psychology*, 16, Article 1730953. <https://doi.org/10.3389/fpsyg.2025.1730953>
- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268.
- Delbio, A. (2025). Reimagining English language teaching: The role of technology integration in 21st century ELT classrooms. *International Journal of Research and Analytical Reviews*, 10(Special Issue 3), 1–8.
- Dimitrov, D. M., & Rumrill, P. D. (2003). Pretest-posttest designs and measurement of change. *Work*, 20(2), 159–165.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59–109.
- Godwin-Jones, R. (2011). Emerging technologies: Mobile apps for language learning. *Language Learning & Technology*, 15(2), 2–11.
- Hermansah, B., Setywati, H., Nasuka, N., & Setiawaty, E. (2025). Enhancing digital competencies through technology integration in vocational education. *Jurnal MENTARI: Manajemen, Pendidikan dan Teknologi Informasi*, 4(1), 40–51. <https://doi.org/10.33050/mentari.v4i1.906>
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125–132.
- JASP. (2026). JASP [Computer software]. <https://jasp-stats.org>
- Kurniawan, A., Syafrizal, & Samanhudi, U. (2026). Development of SpeakUp digital platform based on digital literacy and local wisdom to improve English speaking competence of vocational students. *Multidisciplinary Science Journal*, 9(1), Article 2027087.
- Kurniawati, A., & Lintang Sari, A. (2024). Reducing students' anxiety in speaking English: Teacher's strategies. *Journal of Nusantara Education*, 4(1), 70–77. <https://doi.org/10.57176/jn.v4i1.128>
- Lambert, C., Philp, J., & Nakamura, S. (2017). Learner-generated content and engagement in second language task performance. *Language Teaching Research*, 21(6), 665–680.
- MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning*, 44(2), 283–305.
- Meijers, F. (1998). The development of a career identity. *International Journal for the Advancement of Counselling*, 20(3), 191–207. <https://doi.org/10.1023/A:1005399417256>
- Mentari, M., Abid, A., & Bay, I. W. (2022). Unveiling the sources of English speaking anxiety of a vocational school student: A narrative study. *Journal of English Language Teaching and Linguistics*, 5(2), 211–224. <https://doi.org/10.37905/jetl.v5i2.30245>
- Mirzaei, A., Shafiee Rad, H., & Rahimi, E. (2022). Integrating the ARCS motivational model and flipped teaching in L2 classrooms: A case of EFL expository writing. *Computer Assisted Language Learning*. Advance online publication. <https://doi.org/10.1080/09588221.2022.2068614>
- Muhria, L., Syafrizal, S., & Samanhudi, U. (2023). Blended learning for enhancing EFL speaking performance in Indonesian vocational education. *Journal of Language Teaching and Research*, 14(4), 998–1008.
- Nguyen, T. D., Cannata, M., & Miller, J. (2018). Understanding student behavioral engagement: Importance of student interaction with peers and teachers. *The Journal of Educational Research*, 111(2), 163–174.

- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.
- Ölmezer-Öztürk, E., & Öztürk, G. (2021). Reducing speaking anxiety in EFL classrooms: An explanatory mixed-methods study. *Porta Linguarum*, (36), 249–261. <https://doi.org/10.30827/portalin.v0i36.18018>
- Saputro, T. H., & Hima, A. N. (2024). The impact of data-driven learning on vocational high school students' motivation in English language learning. *AJELP: The Asian Journal of English Language and Pedagogy*, 12(2), 122–138. <https://doi.org/10.37134/ajelp.vol12.2.9.2024>
- Skorikov, V. B., & Vondracek, F. W. (2011). Occupational identity. In S. J. Schwartz, K. Luyckx, & V. L. Vignoles (Eds.), *Handbook of identity theory and research* (pp. 693–714). Springer. https://doi.org/10.1007/978-1-4419-7988-9_29
- Stockwell, G. (2010). Using mobile phones for vocabulary activities: Examining the effect of the platform. *Language Learning & Technology*, 14(2), 95–110.
- Super, D. E. (1980). A life-span, life-space approach to career development. *Journal of Vocational Behavior*, 16(3), 282–298. [https://doi.org/10.1016/0001-8791\(80\)90056-1](https://doi.org/10.1016/0001-8791(80)90056-1)
- Tan, J. P.-L., & Nie, Y. (2015). The role of authentic tasks in promoting twenty-first century learning dispositions. In Y. H. Cho, I. S. Caleon, & M. Kapur (Eds.), *Authentic problem solving and learning in the 21st century* (pp. 19–39). Springer.
- Vickers, A. J., & Altman, D. G. (2001). Statistics notes: Analysing controlled trials with baseline and follow up measurements. *BMJ*, 323(7321), 1123–1124.
- Wang, Y., & Liu, H. (2022). The mediating roles of buoyancy and boredom in the relationship between autonomous motivation and engagement among Chinese senior high school EFL learners. *Frontiers in Psychology*, 13, Article 992279. <https://doi.org/10.3389/fpsyg.2022.992279>
- Wekerle, C., Daumiller, M., & Kollar, I. (2020). Using digital technology to promote higher education learning: The importance of different learning activities and their relations to learning outcomes. *Journal of Research on Technology in Education*, 54(1), 1–17.