

Murat S.Zh.¹, *Baigunisova G.I.², Turusheva L.³

^{1,2} L.N. Gumilyov Eurasian National University

³ EKA University of Applied Sciences

^{1,2} Kazakhstan, Astana

³ Latvia, Riga

¹ ORCID 0009-0001-2166-5242

² ORCID 0000-0001-8964-853X

³ ORCID 0000-0003-4134-7038

* baigunisova@mail.ru

FROM LEARNING STYLES TO EVIDENCE-BASED DIAGNOSTICS: PSYCHOMETRICS OF INDIVIDUAL DIFFERENCES IN ELT

Abstract

For decades, the concept of learning styles has influenced English language teaching (ELT), encouraging teachers to adapt instruction to assumed visual, auditory, or kinesthetic learner preferences. However, this approach has been increasingly criticized for its weak empirical foundation and limited pedagogical value. The purpose of this article is to justify a shift from learning styles to evidence-based psychometric diagnostics of individual differences in ELT. The study focuses on measurable learner characteristics such as L2 motivation, speaking anxiety, self-efficacy, and self-regulation. Methodologically, the article combines theoretical analysis with pilot data from an 8-week school-based quasi-experimental study conducted in Astana, Kazakhstan, where psychometric profiling was integrated into English language instruction. The results showed acceptable reliability of the main diagnostic scales, comparable baseline profiles between experimental and control groups, and stronger learning gains in the experimental condition, particularly in speaking performance. In addition, learners who received psychometrically informed instruction reported more supportive assessment experiences, greater willingness to communicate, and lower anxiety during assessed speaking tasks. These findings are significant because they demonstrate that psychometric diagnostics can provide teachers with more valid and actionable information than traditional learning-style categories. The article concludes that evidence-based diagnostics offer a more scientifically grounded and pedagogically useful approach to individual differences in ELT, improving instructional fit and supporting positive washback in classroom assessment.

Keywords: learning styles, psychometrics, individual differences, ELT, speaking anxiety, washback, evidence-based psychometric diagnostics

Introduction. In contemporary English Language Teaching (ELT), increasing attention is being paid to individual differences among learners. Students in the same classroom may differ significantly in their motivation, language anxiety, self-confidence, self-regulation, willingness to communicate, and classroom participation. These differences influence how learners respond to instructional tasks, interact with peers, participate in speaking activities, and make progress in language learning. Therefore, effective ELT requires not only well-designed teaching materials and communicative tasks, but also a systematic understanding of the psychological and learning-related factors that shape students' performance.

For many years, the concept of learning styles has been widely used in educational practice. According to this approach, learners are often classified as visual, auditory, kinesthetic, or other types of learners. This approach has been attractive to teachers because it seems simple, practical, and easy to apply in the classroom. However, the scientific validity of learning styles has been seriously questioned. Pashler, McDaniel, Rohrer, and Bjork argue that there is insufficient empirical evidence to prove that matching instruction to a learner's preferred learning style improves learning outcomes [1]. In the context of ELT, this criticism is especially relevant because language learning is a complex process that depends on cognitive, affective, social, and behavioural factors rather than on one preferred channel of receiving information.

Recent research in second language acquisition provides a stronger basis for understanding learner differences. Dörnyei and Ryan emphasize that language learners differ in a wide range of psychological and behavioural characteristics, including motivation, anxiety, aptitude, personality, and learning strategies [2]. Similarly, Li, Hiver, and Papi argue that individual differences in second language acquisition should be understood as multidimensional and dynamic factors that influence learners' engagement, performance, and development [3]. These studies show that learner diversity

cannot be fully explained through broad learning-style categories. Instead, it should be examined through evidence-based constructs that can be measured, interpreted, and connected with pedagogical decisions.

Among these constructs, language anxiety is especially important in ELT because it directly affects learners' participation in speaking activities. Papi and Khajavy note that second language anxiety is a complex construct that influences learners' emotional state, classroom behaviour, and language performance [4]. Alrabai also demonstrates that anxiety and motivation play a predictive role in second language (L2) proficiency, showing that affective factors are closely related to language achievement [5]. This is particularly relevant for speaking tasks, where fear of mistakes, public evaluation, and low confidence may prevent learners from using the target language actively.

At the same time, the use of psychometric diagnostics in ELT requires careful attention to the quality of measurement. Sudina emphasizes that second language survey research must consider scale quality, reliability, and construct validity, especially when measuring anxiety and motivation [6]. This means that psychometric tools should not be used mechanically or superficially. They should be selected and interpreted responsibly, with attention to whether they provide meaningful information for teaching and learning.

The issue of validity is central to this discussion. Kane argues that test scores should be validated not only as numerical results, but also in terms of their interpretations and uses [7]. Similarly, Messick's concept of validity and washback highlights that assessment should be evaluated by its educational consequences as well as by its measurement quality [8]. From this perspective, psychometric diagnostics can be useful in ELT only when the results are used ethically and formatively: not for labelling or ranking students, but for adapting instruction, reducing barriers to participation, and supporting learner progress.

Despite this theoretical foundation, in many school contexts, including ELT classrooms in Kazakhstan, individual differences are still often addressed intuitively rather than systematically. Teachers may notice that some learners are passive, anxious, unmotivated, or reluctant to speak, but these observations are not always supported by diagnostic tools or translated into specific pedagogical strategies. This creates a practical problem: teachers recognize learner diversity, but they may lack an evidence-based framework for responding to it.

Therefore, the main problem addressed in this study is the limited use of psychometric diagnostics as a practical tool for adapting ELT instruction to learners' individual needs. The aim of the study is to examine how evidence-based psychometric diagnostics of individual differences can be used as an alternative to the learning styles approach in English language teaching. More specifically, the study aims to develop and test a Psycho-Diagnostic ELT Model that links diagnostic information about learners with concrete pedagogical decisions in the classroom.

The study was conducted with Grade 9 students in a school-based ELT context. The learners were divided into an experimental group (EG) and a control group (CG). In the experimental group, teaching was adapted according to diagnostic data on learners' motivation, language anxiety, self-efficacy, and self-regulation. In the control group, instruction followed the regular teaching approach without systematic psychometric diagnostics. The study examined whether diagnostic-based teaching could contribute to higher motivation, lower language anxiety, more active participation in speaking tasks, and better overall language progress.

The significance of the study lies in providing a practical evidence-based alternative to learning styles in ELT. Instead of labelling students as visual, auditory, or kinesthetic learners, the proposed approach identifies real learning-related needs and uses this information to improve instructional fit, reduce barriers to participation, and support language learning.

Methods and Materials. This study employed a mixed-methods approach within an 8-week school-based quasi-experimental design to examine the effectiveness of evidence-based psychometric diagnostics in English Language Teaching. The quasi-experimental component compared an experimental and a control condition using intact classes, while quantitative and qualitative methods were combined to measure changes in students' English language performance,

motivation, foreign language anxiety, self-efficacy, self-regulation, and classroom participation. The research was conducted in three stages.

Stage 1. Diagnostic and Analytical Stage

At the initial stage, a diagnostic assessment of learners' individual differences was conducted in order to identify baseline levels of L2 motivation, foreign language anxiety, self-efficacy, and self-regulation. In parallel, a theoretical analysis of contemporary approaches to learner individualization in ELT was carried out through a review of research on learning styles, psychometrics, validity, washback, and individual differences in second language acquisition.

This stage provided the conceptual foundation for designing the Psycho-Diagnostic ELT Model, which was developed as an evidence-based alternative to the traditional learning styles approach. The model was based on the idea that teachers should not classify learners as visual, auditory, or kinesthetic types, but should identify measurable learning-related factors and use them for concrete pedagogical decisions.

Stage 2. Experimental Stage

During the experimental stage, the Psycho-Diagnostic ELT Model was implemented in the educational process. The intervention was conducted over an eight-week period in Grade 9 English classes at Nurorda School-Lyceum.

Two Grade 9 classes, 9A and 9C, served as the experimental condition and received psychometrics-informed instruction guided by explicit decision rules. Two Grade 9 classes, 9B and 9D, served as the control condition and followed the regular ELT program without psychometric profiling or systematic psycho-informed adaptations.

The instructional activities in the experimental group included gradual speaking tasks, pair and group rehearsal, delayed or private feedback, micro-goals, learning checklists, progress visualization, structured turn-taking, and choice-based communicative tasks. These activities were selected according to the learners' diagnostic profiles.

The content and functional structure of the proposed model, including its main components, implementation forms, and expected outcomes, are presented in Table 1.

Table 1. Structure and components of the Psycho-Diagnostic ELT Model

Components of the model	Function	Implementation forms	Expected outcomes
Diagnostic component	Identification of learners' individual differences	Psychometric battery, pre-test, teacher observation	Learner profiles based on motivation, anxiety, self-efficacy, and self-regulation
Motivational component	Increasing learners' interest and engagement	Choice-based tasks, real-life topics, progress visualization	Higher motivation and more active participation
Anxiety-reduction component	Reducing fear of speaking and evaluation	Speaking ladder, pair rehearsal, delayed/private feedback	Lower speaking anxiety and greater willingness to communicate
Self-efficacy component	Developing learners' confidence in English tasks	Gradual task difficulty, success-oriented tasks, supportive feedback	Improved confidence and persistence
Self-regulation component	Supporting planning and independent learning	Micro-goals, checklists, short deadlines, formative checkpoints	Better task completion and learning organization

Communicative component	Increasing oral participation	Structured turn-taking, pair work, small-group interaction	Improved speaking performance and classroom interaction
Assessment and reflection component	Monitoring learning outcomes and washback	Post-test, perception questionnaire, learner reflection	Evidence of language progress, positive feedback, and improved learner awareness

Source: developed by the authors.

The model connects diagnostic results with practical classroom actions. Its main function is to help the teacher adapt motivation support, speaking tasks, feedback, and self-regulation strategies according to learners' profiles.

Stage 3. Evaluation Stage

The final stage focused on evaluating the effectiveness of the proposed model through comparison of pre- and post-intervention results between the experimental and control groups. At this stage, students completed a language post-test, a repeated psychometric questionnaire, and a brief washback/perception questionnaire. Teacher observations and short learner reflections were also collected to identify changes in participation, confidence, and attitudes toward English lessons.

The evaluation stage made it possible to compare both achievement-related outcomes and psychological indicators. Particular attention was paid to speaking performance, because oral participation is strongly influenced by anxiety, self-confidence, and motivation.

Participants. The study was conducted at Nurorda School-Lyceum in Astana, Kazakhstan. A total of 40 Grade 9 students participated in the research. The participants were selected from four intact classes: 9A, 9B, 9C, and 9D, with 10 students in each class.

Data Collection Instruments. Several data collection tools were used to ensure methodological triangulation. First, a psychometric questionnaire measured learners' individual differences across four constructs, each assessed on a five-point Likert scale. L2 motivation was measured using the L2 Motivation Scale, a researcher-developed instrument capturing goal orientation and effort persistence (6 items, $\alpha = .84$). Speaking and evaluation anxiety was measured using the Speaking/Evaluation Anxiety Scale, a short-form instrument aligned with the Foreign Language Classroom Anxiety framework (6 items, $\alpha = .88$). L2 self-efficacy was measured using the L2 Self-Efficacy Scale, a researcher-developed instrument of task-specific items reflecting confidence in classroom English performance, informed by Bandura's self-efficacy framework (5 items, $\alpha = .82$). Self-regulation was measured using the Self-Regulation Scale, a researcher-developed instrument of brief indicators assessing planning, monitoring, and persistence suitable for school contexts, informed by Zimmerman's model of self-regulated learning (5 items, $\alpha = .79$). Composite scores for each scale were computed as the mean of item responses, with reverse-coded items recalculated prior to averaging. Scale reliability was assessed using Cronbach's alpha as a measure of internal consistency; all coefficients exceeded the conventional threshold of .70, indicating acceptable reliability for classroom-based research.

Instrument Adaptation and Language of Administration. The Speaking/ Evaluation Anxiety Scale was adapted from the Foreign Language Classroom Anxiety framework; item wording was shortened and simplified for Grade 9 readability while preserving the original construct definitions. The remaining scales (L2 Motivation, L2 Self-Efficacy, and Self-Regulation) were researcher-developed for this study and were not adaptations of existing published instruments and therefore did not require a formal translation-adaptation procedure. All instruments, including the language-achievement test and the washback/perception questionnaire, were administered in English, consistent with the communicative focus of the intervention and students' Grade 9 proficiency level, which was sufficient for comprehending the simplified item wording.

Second, pre-test and post-test tasks were used to assess English language achievement before and after the intervention. The test consisted of three components with a maximum total score of 100 points: a combined reading and listening component (50 points), a lexico-grammatical component

(30 points), and a speaking component (20 points). Parallel test forms of equivalent difficulty were used at pre-test and post-test to support comparability across measurement points. The total score was calculated as the sum of the three component scores. Oral performance was evaluated using a speaking rubric based on four criteria: accuracy, fluency, lexical range, and interaction, each contributing proportionally to the 20-point speaking score.

The speaking rubric evaluated oral performance according to four analytic criteria, each scored on a 0-5 point scale, for a maximum total of 20 points: (a) accuracy – grammatical and lexical correctness with minimal errors that do not impede communication; (b) fluency – smoothness and pace of speech with limited hesitation or unnatural pauses; (c) lexical range – variety and appropriateness of vocabulary relative to the task; and (d) interaction – ability to initiate, sustain, and respond appropriately within a spoken exchange. Each criterion was scored independently by the rater using descriptive band levels (0 = minimal/no evidence, 1-2 = limited, 3 = adequate, 4 = good, 5 = excellent), and the final speaking score was calculated as the sum of the four criterion scores.

In addition, teacher observation notes were collected during the intervention to record students' participation, confidence, speaking initiative, task completion, and cooperation. Classroom observation was conducted using a structured observation checklist completed by the teacher-researcher during three lessons per week across the eight-week intervention period (24 lessons in total for each condition). For each of the five indicators reported in Table 5 (active participation in pair work, speaking initiative, confidence during speaking, task completion, and cooperation with classmates), the observer recorded, for every student, whether the target behaviour was demonstrated during each observed lesson (1 = demonstrated, 0 = not demonstrated). Active participation in pair work was recorded when a student engaged in the assigned pair or group task without prompting; speaking initiative was recorded when a student volunteered to speak without being called on; confidence during speaking was recorded based on fluency and absence of visible hesitation or avoidance behaviours; task completion was recorded when a student finished the assigned classroom task within the given time; and cooperation with classmates was recorded when a student assisted or responded constructively to peers during group work. The percentage reported for each indicator in Table 5 was calculated as the number of lessons in which the behaviour was observed for a given student, averaged across all students in the condition, divided by the total number of observed lessons, and expressed as a percentage. At the end of the experiment, a washback/perception questionnaire was administered to identify students' attitudes toward learning, perceived support, anxiety, confidence, and willingness to participate in English lessons.

The psychometric battery and perception questionnaire were administered through Google Forms, and the collected data were organized in Microsoft Excel.

Data Analysis. Quantitative data were analyzed using descriptive statistics. Mean scores and pre-test/post-test gains were calculated for the experimental and control groups to identify changes in English achievement and speaking performance. Questionnaire data were analyzed by comparing average scores for motivation, anxiety, self-efficacy, and self-regulation. For the primary outcome indicators (general achievement, speaking performance, and anxiety during assessed speaking), between-group differences at posttest were additionally tested using independent-samples Welch's t-tests, with Cohen's d calculated to estimate effect size (see Table 6). Qualitative data from teacher observations and learner reflections were summarized thematically, focusing on participation, confidence, cooperation, and attitudes toward feedback.

Ethical Considerations. Since the participants were minors, the study followed standard ethical principles of classroom-based research. School-level approval was obtained, and the procedures were coordinated with the administration. Parental consent and student assent were secured before data collection.

Participation was voluntary and unrelated to students' grades. All data were anonymized and used only for research and pedagogical purposes. Psychometric results were not used for ranking, labelling, tracking, or punitive decisions. Their only function was to support instruction and help the teacher respond more effectively to learners' needs.

Results and Discussion. The results of the study show that psychometrics-informed instruction had a positive influence on Grade 9 students' English language achievement, speaking performance, motivation, self-efficacy, self-regulation, and classroom participation. The experimental group, consisting of classes 9A and 9C, demonstrated greater improvement than the control group, consisting of classes 9B and 9D. This confirms the main assumption of the study: individualization in ELT should be based not on fixed learning styles, but on measurable learner characteristics that can guide practical teaching decisions.

At the initial stage of the experiment, both groups had relatively similar results. This means that the experimental and control conditions were comparable before the intervention. After the implementation of the Psycho-Diagnostic ELT Model, however, the experimental group showed stronger progress in both linguistic and psychological indicators. Table 2 presents the general English achievement scores of both groups before and after the experiment.

Table 2. General English achievement before and after the experiment

Group	Number of students	Pre-test average score	Post-test average score	Growth
Experimental group, 9A and 9C	20	24.2	32.1	+7.9
Control group, 9B and 9D	20	24.5	28.6	+4.1

Source: developed by the authors.

The data in Table 2 show that both groups improved by the end of the experiment; however, the experimental group demonstrated greater progress (+7.9) than the control group (+4.1). This suggests that psychometrics-informed instruction created more favourable conditions for language learning by adapting teaching decisions to learners' diagnostic profiles, including speaking anxiety, self-regulation, and motivation. This magnitude of improvement is broadly consistent with international quasi-experimental evidence on assessment-informed instruction in EFL contexts. Zhang, Quan, Zhang, and Zhang conducted an eight-week quasi-experimental intervention with middle school EFL students that closely parallels the design of the present study and found that formative-assessment-based instruction produced greater gains in both achievement and motivation compared with conventional teaching [9]. This close correspondence in sample age, intervention length, and design supports the plausibility of the achievement gain observed here, although the effect should still be interpreted cautiously given the small, single-site sample. A graphical comparison of pre-test and post-test scores is shown in Figure 1 below.

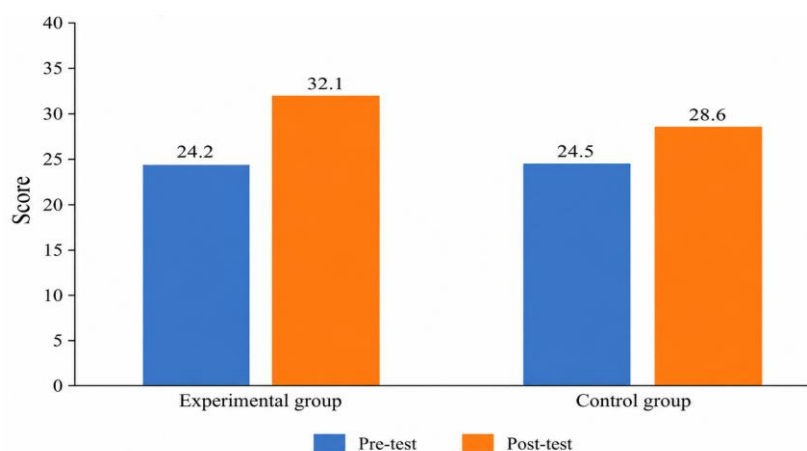


Figure 1. Dynamics of English achievement before and after the experiment

Source: developed by the authors.

Figure 1 illustrates that the experimental group demonstrated a more visible increase in general English achievement. This result supports the argument that diagnostic-based teaching may be more effective than general instruction when learners' psychological and educational needs are systematically considered.

Speaking performance was analyzed separately because oral communication is one of the most sensitive areas in ELT. Learners may possess adequate lexical and grammatical knowledge yet consistently avoid oral participation, as speaking anxiety, fear of error, and low self-confidence constitute significant affective barriers to language production. A detailed comparison of speaking performance scores is provided in Table 3.

Table 3. Speaking performance before and after the experiment

Group	Pre-test speaking score	Post-test speaking score	Growth
Experimental group	5.7	7.8	+2.1
Control group	5.8	6.7	+0.9

Source: developed by the authors.

As shown in Table 3, the experimental group improved its speaking score from 5.7 to 7.8, while the control group increased from 5.8 to 6.7. The higher growth in the experimental group (+2.1 compared with +0.9) can be explained by psycho-informed speaking strategies, including a speaking ladder, pair rehearsal, gradual public performance, and delayed/private feedback, which reduced fear of mistakes and supported learners' confidence. This finding is consistent with international evidence that structured, anxiety-sensitive speaking support disproportionately benefits oral performance relative to other language skills. Jin, Zhang, and MacIntyre found that structured speaking-commitment interventions reduced foreign language classroom anxiety and improved learners' confidence and engagement [10]. A comparable pattern was reported in a Chinese EFL context, where a speaking-support intervention – delivered through an AI-speaking assistant rather than a teacher-led format – produced a significant reduction in foreign language anxiety and a significant increase in willingness to communicate in the experimental group, with no such change in the control group [11]. Although the delivery mechanism differs from the teacher-mediated approach used in the present study, the convergence in outcome pattern (EG improvement vs. CG stability) suggests that reducing evaluative pressure around speaking, regardless of the specific method, reliably increases participation and confidence. Table 4 outlines the corresponding changes in motivation, anxiety, self-efficacy, and self-regulation.

Table 4. Changes in motivation, anxiety, self-efficacy, and self-regulation

Indicator	Experimental group before	Experimental group after	Control group before	Control group after
L2 motivation	3.4	4.2	3.5	3.8
Foreign language anxiety	3.8	2.9	3.7	3.4
L2 self-efficacy	3.2	4.0	3.3	3.6
Self-regulation	3.1	3.8	3.2	3.5

Source: developed by the authors.

As shown in Table 4, both groups demonstrated positive changes in psychological indicators, but the experimental group showed stronger improvement. Motivation increased from 3.4 to 4.2, self-

efficacy from 3.2 to 4.0, and self-regulation from 3.1 to 3.8, while foreign language anxiety decreased from 3.8 to 2.9. These results suggest that psychometrics-informed instruction supported not only language achievement, but also learners' confidence, emotional safety, and ability to organize their own learning. The observed decrease in speaking anxiety alongside increased participation is consistent with the broader intervention literature on foreign language anxiety. Alrabai's causal model established a predictive link between anxiety, motivation, and L2 proficiency [5, p. 3127]; the present findings extend this by showing the same pattern under a classroom intervention rather than a correlational design. This is further supported by direct experimental evidence: in a randomized study, Jin, Dewaele, and MacIntyre found that a structured anxiety-reduction intervention produced a significant decrease in foreign language classroom anxiety in the experimental group, while anxiety remained stable in the control group [12] – the same divergent pattern observed between the experimental and control classes in Table 4. This supports Papi and Khajavy's account of anxiety as a dynamic, emotionally mediated construct that can be shifted through targeted classroom intervention rather than a static trait [4, p. 138]. The improvement in self-efficacy (3.2 to 4.0) is consistent with Li's quasi-experimental study of a self-efficacy-based instructional intervention, which found that structured, confidence-building instruction reliably raised learners' self-efficacy beliefs and, in turn, supported their motivation and engagement in EFL learning [13]. Similarly, the improvement in self-regulation (3.1 to 3.8) parallels the broader finding that structured, formative, decision-rule-based instruction supports learners' self-organizational skills alongside achievement gains [9, p. 11]. These parallels suggest that the gains observed in this study reflect a broader, internationally replicated pattern linking diagnostic, learner-responsive instruction to stronger self-regulatory and motivational outcomes, rather than being idiosyncratic to the Kazakhstani context.

Classroom observation data also supported the quantitative results. The teacher observed students' participation, speaking initiative, confidence, task completion, and cooperation with classmates. These results are presented in Table 5.

Table 5. Classroom observation results after the experiment

Criterion	Experimental group	Control group
Active participation in pair work	85%	65%
Speaking initiative	78%	56%
Confidence during speaking	74%	58%
Task completion	88%	76%
Cooperation with classmates	82%	70%

Source: developed by the authors.

The observation results show that students in the experimental group became more active and confident during English lessons. The largest difference was found in speaking initiative: 78% in the experimental group compared with 56% in the control group. This means that students in the experimental classes were more willing to start speaking, answer questions, and participate in pair or group discussions. This gap mirrors behavioural findings from classroom-based anxiety-reduction studies internationally. In the Chinese EFL intervention discussed above, structured speaking support produced a comparable divergence in classroom participation between experimental and control conditions [11, p. 18]. This convergence across different educational contexts and delivery mechanisms suggests that the participation gains found here are unlikely to be a context-specific artifact and instead reflect a general mechanism: reducing evaluative threat around speaking reliably increases learners' willingness to initiate spoken interaction [10, p. 6; 12, p. 24].

The difference can be explained by the structure of the Psycho-Diagnostic ELT Model. Students with low participation were supported through clear group roles, structured turn-taking, and smaller-group interaction. These strategies reduced uncertainty and helped passive students become more involved in the lesson. Table 6 consolidates all key indicators from the study.

Table 6. Summary of main results

Indicator	Experimental group growth	Control group growth	t (df)	p	d
General English achievement	+7.9	+4.1	t(38.00) = 1.33	.190	0.42
Speaking performance	+2.1	+0.9	t(37.94) = 1.98	.055	0.63
Motivation	+0.8	+0.3	—	—	—
Anxiety reduction	-0.9	-0.3	t(37.9) = -2.14	.039	-0.68
Self-efficacy	+0.8	+0.3	—	—	—
Self-regulation	+0.7	+0.3	—	—	—

Source: developed by the authors.

Note: t, p, and d were calculated for the primary outcome indicators only (see Data Analysis). Motivation, self-efficacy, and self-regulation were analyzed descriptively due to the exploratory scope of this pilot. The statistical indicators are defined as follows: t refers to the t-test statistic used to assess differences between groups (with degrees of freedom in parentheses), p indicates the probability value (statistical significance level, conventionally set at $p < .05$), and d (Cohen's d) represents the effect size, reflecting the magnitude of the difference.

The summary of results clearly shows that the experimental group outperformed the control group in all measured indicators. The most important improvements were observed in speaking performance, motivation, anxiety reduction, and general English achievement. Taken together, this pattern – with the experimental group outperforming the control group on every indicator – corresponds closely to what international washback and SLA intervention research would predict: diagnostic-based, anxiety-sensitive instruction tends to produce convergent gains across achievement, affective, and behavioural indicators simultaneously, rather than isolated improvement in a single domain [4, p. 138; 5, p. 3131; 10, p. 7; 12, p. 16]. This convergence strengthens the plausibility of the underlying mechanism proposed by the Psycho-Diagnostic ELT Model, even though the pilot's small sample size means the findings should be treated as preliminary rather than conclusive. In addition to these achievement-related outcomes, washback and perception indicators collected at Week 8 are reported in Table 7.

Table 7. Washback/perception indicators by condition (Week 8)

Indicator (1–5)	Experimental M (SD)	Control M (SD)	t (df)	p	d
Perceived assessment support	4.08 (0.55)	3.46 (0.62)	t(37.5) = 3.35	.002	1.06
Fairness/clarity	4.01 (0.51)	3.58 (0.57)	t(37.5) = 2.51	.016	0.80

Willingness to communicate	3.79 (0.60)	3.25 (0.63)	$t(37.9) = 2.78$	.009	0.88
Anxiety in assessed speaking	2.62 (0.66)	3.08 (0.70)	$t(37.9) = -2.14$	.039	-0.68
Attitude toward feedback	4.15 (0.50)	3.62 (0.60)	$t(36.8) = 3.03$	.004	0.96

Source: developed by the authors.

Note: M = mean; SD = standard deviation. Independent-samples Welch's t-tests were used to compare group means; effect sizes are reported as Cohen's d. t refers to the t-test statistic (with degrees of freedom in parentheses), and p indicates the probability value (statistical significance level, conventionally set at $p < .05$). Higher values indicate more positive perceptions, except for anxiety (lower is better).

As shown in Table 7, students in the experimental group reported significantly more positive assessment experiences than those in the control group across all five indicators. Perceived assessment support ($d = 1.06$) and attitude toward feedback ($d = 0.96$) showed the largest effects, followed by willingness to communicate ($d = 0.88$) and fairness/clarity ($d = 0.80$) – all representing large effect sizes. Anxiety during assessed speaking was significantly lower in the experimental group ($d = -0.68$), consistent with the anxiety-reduction findings reported in Table 4. These results provide quantitative support for the washback claims stated in the Introduction and Abstract: psychometrics-informed instruction was associated not only with stronger achievement and speaking gains, but also with a more supportive and less threatening assessment experience.

These findings support the conceptual logic of the study: psychometric diagnostics → learner profile → pedagogical decision rules → adapted instruction → improved learning outcomes.

The study therefore argues that psychometric diagnostics should not be treated as a separate testing procedure. Its value lies in its connection with teaching practice: the teacher collects diagnostic information to understand what kind of support each learner needs.

From a scientific point of view, the findings contribute to the discussion about learning styles and evidence-based teaching. The learning styles approach usually classifies students as visual, auditory, or kinesthetic learners. However, such classification gives limited practical information to the teacher. For example, knowing that a student prefers visual materials does not explain why the student avoids speaking, feels anxious, or fails to complete independent tasks.

In contrast, psychometric diagnostics provides more actionable data. If a student has high speaking anxiety, the teacher can use gradual speaking exposure. If a student has low motivation, the teacher can provide meaningful choices and real-life tasks. If a student has low self-regulation, the teacher can introduce checklists, micro-goals, and progress monitoring. Thus, psychometric diagnostics helps the teacher move from general assumptions to concrete pedagogical decisions.

The findings also support the idea of positive washback. In traditional assessment, tests often measure only final results. In this study, diagnostic tools influenced the learning process itself. Assessment data were used to adapt teaching, reduce anxiety, increase participation, and improve feedback practices. Therefore, assessment created a positive effect on teaching and learning.

This finding is consistent with Kane's argument that validity should be evaluated through the interpretation and use of scores rather than through psychometric properties alone [7, p. 1]: in this study, diagnostic scores were not treated as endpoints but were explicitly translated into pedagogical decision rules, which is precisely the kind of validated "use" argument Kane's framework requires. It also aligns with foundational washback scholarship showing that test consequences are mediated primarily by teachers, who translate assessment information into instructional adjustment rather than the test format acting alone [14], a claim recently reaffirmed at scale across multiple stakeholder contexts [15]. This dependency is directly reflected in the design of the Psycho-Diagnostic ELT Model, where the teacher, not the diagnostic instrument itself, decides how motivation, anxiety, or self-regulation data are converted into classroom action.

Reliability of the diagnostic scales used in this study was acceptable across constructs. This is a relevant methodological point: Sudina's earlier review of second-language survey research found that reliability reporting in anxiety and motivation research is frequently incomplete, and that item-total correlations are rarely reported even when overall alpha values are provided [6, p. 1173]. A subsequent, larger-scale methodological synthesis focusing specifically on anxiety and willingness-to-communicate scales confirmed this pattern across 232 peer-reviewed articles and 385 scales, finding a general lack of evidence for scale content and construct validity [16]. In line with these recommendations, future iterations of the Psycho-Diagnostic ELT Model should also report item-total correlations and, where feasible, confirmatory factor analysis to establish structural validity rather than internal consistency alone [6, p. 1184; 16, p. 1449].

To verify that the observed between-group differences were not due to chance, independent-samples Welch's t-tests and Cohen's d were calculated for the primary outcome indicators. The difference in speaking performance approached statistical significance ($t(37.94) = 1.98, p = .055$) with a moderate-to-large effect size ($d = 0.63$), while the reduction in anxiety during assessed speaking was statistically significant ($t(37.9) = -2.14, p = .039, d = -0.68$). The difference in general English achievement, although favouring the experimental group, did not reach conventional significance ($t(38.00) = 1.33, p = .190, d = 0.42$), which is consistent with the modest sample size of this pilot study. Statistical testing was conducted for the primary outcome indicators; motivation, self-efficacy, and self-regulation scores were analyzed descriptively due to the exploratory scope of this pilot. These results indicate that the experimental condition produced a moderate practical effect, particularly for speaking-related outcomes, even where statistical significance was only marginal.

At the same time, the results should be interpreted cautiously. The experiment involved a small sample of 40 students from one school, and the modest sample size limited statistical power for detecting significance in several outcomes.

Although the observed pattern of results is consistent with the mechanisms proposed by the Psycho-Diagnostic ELT Model, alternative explanations should also be considered given the quasi-experimental, non-randomized design of this pilot. First, because students were not randomly assigned to conditions, pre-existing differences between the intact classes – beyond the psychometric variables measured at baseline – cannot be fully ruled out as contributing factors. Second, a Hawthorne-type effect is possible: students in the experimental classes may have responded positively simply because they received additional attention and a novel instructional approach, independent of the specific diagnostic content of the intervention. Third, because the same teacher-researcher delivered instruction to the experimental group and was directly involved in data collection, teacher enthusiasm or expectancy effects cannot be entirely separated from the effect of the model itself. Finally, the experimental group's additional exposure to structured speaking practice, regardless of its diagnostic basis, may partly explain the gains in speaking performance, as increased practice opportunities alone are known to improve oral proficiency. Disentangling these alternative explanations from the effect of psychometric diagnostics specifically would require a more controlled design, such as comparing the diagnostic intervention against an active control condition receiving equivalent instructional time and teacher attention without profile-based adaptation.

Nevertheless, the results demonstrate that the Psycho-Diagnostic ELT Model can be used as a feasible classroom tool. It helps teachers understand learners more systematically and respond to their needs more effectively. The study shows that individual differences in ELT should not be reduced to learning styles. They should be studied through measurable constructs such as motivation, anxiety, self-efficacy, and self-regulation.

Thus, the experimental data suggest that psychometrics-informed instruction is associated with gains in students' speaking performance and reduced speaking anxiety in particular, alongside broader positive trends in confidence, motivation, and classroom participation; however, given that the difference in general English achievement did not reach conventional statistical significance, these findings should be interpreted as preliminary evidence rather than confirmed effects.

Conclusion. The results of the study provide preliminary support for psychometrics-informed instruction as a promising evidence-based alternative to the traditional learning styles approach in

English Language Teaching, particularly for speaking-related and anxiety-related outcomes; broader claims of overall effectiveness should await replication with larger samples and stronger significance testing across all indicators. The experimental group, which received instruction based on the Psycho-Diagnostic ELT Model, demonstrated stronger progress than the control group. In particular, general English achievement increased from 24.2 to 32.1 in the experimental group, while the control group improved from 24.5 to 28.6. Speaking performance also improved more noticeably in the experimental group, rising from 5.7 to 7.8, compared with the control group's increase from 5.8 to 6.7. Positive changes were also observed in motivation, self-efficacy, self-regulation, and classroom participation, while foreign language anxiety decreased.

The scientific novelty of the study lies in the development and pilot implementation of the Psycho-Diagnostic ELT Model, which differs from existing approaches to learner individualization in three specific ways. First, unlike the learning styles approach discussed above, the model does not classify students into fixed learner types; instead, it relies on measurable, empirically grounded constructs. Second, unlike much of the existing individual-differences scholarship in SLA, which remains primarily descriptive – identifying correlations between motivation, anxiety, or self-regulation and learning outcomes without specifying how teachers should act on this information – the model is operational: it introduces an explicit construct-to-decision mapping (Table 1) that translates each diagnostic category (low/medium/high) into a defined instructional response (e.g., staged speaking exposure for high anxiety, micro-goals for low self-regulation). Third, unlike general formative assessment or differentiated instruction frameworks, which adjust teaching based on academic performance alone, the model integrates affective and motivational diagnostics into the same decision cycle as achievement data and explicitly embeds washback monitoring as part of the model's validity evidence rather than as a separate, optional evaluation step. The authorial contribution of this study is therefore not any single diagnostic construct in isolation – each of which has established roots in prior research – but their integration into a single, minimal, classroom-feasible diagnostic battery bundled with an explicit instructional decision logic: diagnosis → learner profile → pedagogical decision rules → adapted instruction → improved learning outcomes.

The practical value of the study is that the proposed model can be applied in real school ELT classrooms. It helps teachers identify learners with high speaking anxiety, low motivation, weak self-regulation, or low self-efficacy and select appropriate teaching strategies, such as gradual speaking exposure, pair rehearsal, delayed feedback, micro-goals, checklists, choice-based tasks, and progress visualization. Thus, psychometric diagnostics functions not as an additional testing burden, but as a formative tool for improving instruction and creating positive washback.

However, the study has several limitations related to its pilot scope: a single-site, intact-class design constrains generalizability, and longer-term follow-up was not possible within the eight-week window. Future research should replicate the model across multiple schools, extend the intervention period, and test its applicability across different age groups and language proficiency levels.

In conclusion, the findings suggest that learner individualization in ELT should move beyond unsupported learning styles and focus on measurable, pedagogically meaningful individual differences. The Psycho-Diagnostic ELT Model can serve as a practical evidence-based framework for improving students' motivation, confidence, speaking participation, and overall language achievement.

REFERENCES

- 1 Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest*, 9(3), 105-119. DOI: <https://doi.org/10.1111/j.1539-6053.2009.01038.x>
- 2 Dornyei, Z., & Ryan, S. (2015). *The Psychology of the Language Learner Revisited* (1st ed.). Routledge, 274 p. DOI: <https://doi.org/10.4324/9781315779553>
- 3 Li, S., Hiver, P., Papi, M. (2022). Individual differences in second language acquisition: Theory, research, and practice. In: *The Routledge Handbook of Second Language Acquisition and Individual Differences* (1st ed.). Routledge, 32. DOI: <https://doi.org/10.4324/9781003270546-2>

- 4 Papi, M., & Khajavy, G. H. (2023). Second language anxiety: Construct, effects, and sources. *Annual Review of Applied Linguistics*, 43, 127-139. DOI: <https://doi.org/10.1017/S0267190523000028>
- 5 Alrabai, F. (2022). The predictive role of anxiety and motivation in L2 proficiency: An empirical causal model. *Language Teaching Research*, 29(7), 3104-3142. DOI: <https://doi.org/10.1177/13621688221136247>
- 6 Sudina, E. (2021). Study and scale quality in second language survey research, 2009–2019: The case of anxiety and motivation. *Language Learning*, 71, 1149-1193. DOI: <https://doi.org/10.1111/lang.12468>
- 7 Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50, 1–73. DOI: <https://doi.org/10.1111/jedm.12000>
- 8 Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241-256. DOI: <https://doi.org/10.1177/026553229601300302>
- 9 Zhang, H., Quan, J., Zhang, N., & Zhang, L. (2025). Investigating the effects of formative assessment on EFL students' achievement and motivation: a Self-Determination Theory perspective. *Frontiers in Psychology*, 16, 1664871. DOI: <https://doi.org/10.3389/fpsyg.2025.1664871>
- 10 Jin, Y., Zhang, L.J. and MacIntyre, P.D. (2020). Contracting students for the reduction of foreign language classroom anxiety: An approach nurturing positive mindsets and behaviors. *Frontiers in Psychology*, 11, 1471. DOI: <https://doi.org/10.3389/fpsyg.2020.01471>
- 11 Zhang, C., Meng, Y., & Ma, X. (2024). Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. *System*, 121, 103259. DOI: <https://doi.org/10.1016/j.system.2024.103259>
- 12 Jin, Y., Dewaele, J.-M., & MacIntyre, P. D. (2021). Reducing anxiety in the foreign language classroom: A positive psychology approach. *System*, 101, 102604. DOI: <https://doi.org/10.1016/j.system.2021.102604>
- 13 Li, J., Wang, C., Zhao, Y., Li, Y. (2023). Boosting learners' confidence in learning English: Can self-efficacy-based intervention make a difference? *TESOL Quarterly*, 58(4), 1518-1547. DOI: <https://doi.org/10.1002/tesq.3292>
- 14 Rahman, K. A., Rukanuddin, M., Chowdhury, M. S. Y., Ahmed, S., & Seraj, P. M. I. (2023). Recognizing stakeholders and factors mediating washback in language testing. *Education Research International*, 1, Article 5548723. DOI: <https://doi.org/10.1155/2023/5548723>
- 15 Watanabe, Y. (2004). Teacher factors mediating washback. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in Language Testing: Research Contexts and Methods* (pp. 129-146). Routledge. DOI: <https://doi.org/10.4324/9781410609731-17>
- 16 Sudina, E. (2023). Scale quality in second-language anxiety and WTC: A methodological synthesis. *Studies in Second Language Acquisition*, 45(5), 1427-1455. DOI: <https://doi.org/10.1017/S02722263122000560>

Мұрат С.Ж.¹, *Байгунисова Г.И.², Турушева Л.³

^{1,2} Л.Н. Гумилев атындағы Еуразия ұлттық университеті

³ ЕКА қолданбалы ғылымдар университеті

^{1,2} Қазақстан, Астана

³ Латвия, Рига

ОҚУ СТИЛЬДЕРІ ТҰЖЫРЫМЫНАН ДӘЛЕЛДІ ДИАГНОСТИКАҒА БЕТБҰРЫС: АҒЫЛШЫН ТІЛІН ОҚЫТУДАҒЫ ЖЕКЕ ЕРЕКШЕЛІКТЕРДІ ПСИХОМЕТРИЯЛЫҚ ТАЛДАУ

Аңдатпа

Ондаған жылдар бойы оқу стильдері тұжырымдамасы ағылшын тілін оқытуға ықпал етіп, оқытушыларды білім алушылардың болжамды визуалды, аудиалды немесе кинестетикалық қалауларына сәйкес оқыту үдерісін бейімдеуге ынталандырып келді. Алайда соңғы жылдары бұл тәсіл эмпирикалық негізінің әлсіздігі және педагогикалық құндылығының шектеулілігі себепті жиі сынға алынуда. Осы мақаланың мақсаты — оқу стильдері тұжырымдамасынан ағылшын тілін оқытудағы жеке айырмашылықтарды дәлелді психометриялық

диагностикалауға көшудің қажеттілігін негіздеу. Зерттеу екінші тілді меңгеруге деген мотивация, сөйлеу кезіндегі мазасыздық, өзіндік тиімділік және өзін-өзі реттеу сияқты білім алушылардың өлшенетін сипаттамаларына бағытталған.

Әдіснамалық тұрғыдан мақала теориялық талдауды Қазақстанның Астана қаласындағы мектеп базасында жүргізілген 8 апталық квазиэксперименттік зерттеудің пилоттық деректерімен ұштастырады. Бұл зерттеу аясында психометриялық профильдеу ағылшын тілін оқыту үдерісіне енгізілді. Нәтижелер негізгі диагностикалық шкалалардың сенімділігі қолайлы деңгейде екенін, эксперименттік және бақылау топтарының бастапқы профильдері өзара салыстырмалы болғанын, сондай-ақ эксперименттік жағдайда оқу нәтижелерінің, әсіресе ауызша сөйлеу көрсеткіштері бойынша, анағұрлым айқын артқанын көрсетті. Сонымен қатар психометриялық деректер ескеріле отырып оқытылған білім алушылар бағалаудың қолдаушы сипатта болғанын, коммуникацияға түсуге дайындығының артқанын және бағаланатын ауызша тапсырмаларды орындау кезінде мазасыздық деңгейінің төмендегенін атап өтті.

Алынған нәтижелердің маңыздылығы мынада: олар психометриялық диагностиканың оқытушыларға дәстүрлі оқу стильдері категорияларына қарағанда анағұрлым валидті әрі тәжірибеде қолдануға жарамды ақпарат бере алатынын көрсетеді. Мақалада дәлелді диагностика ағылшын тілін оқытудағы жеке айырмашылықтарды ескерудің ғылыми тұрғыдан неғұрлым негізделген әрі педагогикалық тұрғыдан тиімді тәсілін ұсынады деген қорытынды жасалады. Бұл тәсіл оқытуды білім алушылардың ерекшеліктеріне сәйкестендіруді күшейтіп, сыныптағы бағалаудың оң кері ықпалын қолдауға мүмкіндік береді.

Түйінді сөздер: оқу стильдері, психометрика, жеке айырмашылықтар, ағылшын тілін оқыту, сөйлеу кезіндегі мазасыздық, бағалаудың кері ықпалы, дәлелді психометриялық диагностика.

Мурат С.Ж.¹, *Байгунисова Г.И.², Турушева Л.³

^{1,2} Евразийский национальный университет имени Л. Н. Гумилёва

³ Университет прикладных наук ЕКА

^{1,2} Казахстан, Астана

³ Латвия, Рига

ОТ УЧЕБНЫХ СТИЛЕЙ К ДОКАЗАТЕЛЬНОЙ ДИАГНОСТИКЕ: ПСИХОМЕТРИКА ИНДИВИДУАЛЬНЫХ РАЗЛИЧИЙ В ПРЕПОДАВАНИИ АНГЛИЙСКОГО ЯЗЫКА

Аннотация

На протяжении десятилетий концепция учебных стилей оказывала влияние на преподавание английского языка, побуждая преподавателей адаптировать обучение к предполагаемым визуальным, аудиальным или кинестетическим предпочтениям учащихся. Однако в последние годы данный подход всё чаще подвергается критике из-за слабой эмпирической обоснованности и ограниченной педагогической ценности. Цель настоящей статьи — обосновать переход от концепции учебных стилей к доказательной психометрической диагностике индивидуальных различий в преподавании английского языка. Исследование сосредоточено на измеряемых характеристиках учащихся, таких как мотивация к изучению второго языка, тревожность при говорении, самозффективность и саморегуляция.

С методологической точки зрения статья сочетает теоретический анализ с пилотными данными 8-недельного квазиэкспериментального исследования, проведенного на базе школы в Астане, Казахстан. В рамках данного исследования психометрическое профилирование было интегрировано в процесс обучения английскому языку. Результаты показали приемлемую надёжность основных диагностических шкал, сопоставимость исходных профилей экспериментальной и контрольной групп, а также более выраженный прирост учебных результатов в экспериментальных условиях, особенно в показателях устной речи. Кроме того, учащиеся, обучавшиеся с учётом психометрических данных, отмечали более поддерживающий характер оценивания, большую готовность к коммуникации и более низкий уровень тревожности во время оцениваемых устных заданий.

Значимость полученных результатов заключается в том, что они демонстрируют: психометрическая диагностика способна предоставить преподавателям более валидную и практически применимую информацию, чем традиционные категории учебных стилей. В статье делается вывод о том, что доказательная диагностика предлагает более научно обоснованный и педагогически полезный подход к учёту индивидуальных различий в преподавании английского языка, способствуя повышению соответствия обучения особенностям учащихся и поддерживая положительный обратный эффект оценивания в классе.

Ключевые слова: учебные стили, психометрика, индивидуальные различия, преподавание английского языка, тревожность при говорении, обратный эффект оценивания, доказательная психометрическая диагностика.

Received: 09.06.2026

Approved after peer review: 17.09.2026

Accepted for publication: 25.09.2026