Quality
Journal of Education, Arabic and Islamic Studies

ISSN 3025-6569
https://quality.pdfaii.or.id
Volume 4. No 4. (2026)

Research Article

Rules In Context: Teaching Sat Writing Via the Ten Essential Rules to Efl Learners

Agung Pranoto Kadiatmaja
Sekolah Tinggi Agama Islam Ali bin Abi Thalib Surabaya
E-mail: agungpranotokadiatmaja@stai-ali.ac.id

Oscar Wardhana Windro Saputro
Sekolah Tinggi Agama Islam Ali bin Abi Thalib Surabaya
E-mail: oscar.wardhana@stai-ali.ac.id

Maryono
Sekolah Tinggi Agama Islam Ali bin Abi Thalib Surabaya
E-mail: maryono003@gmail.com

Copyright © 2026 by Authors, Published by Quality, Journal of Education, Arabic and Islamic Studies. This is an open access article under the CC BY License: (https://creativecommons.org/licenses/by-sa/4.0).

Received : August 28, 2026
Revised : September 19, 2026
Accepted : October 18, 2026
Available online : November 27, 2026

How to Cite: Agung Pranoto Kadiatmaja, Oscar Wardhana Windro Saputro, & Maryono. (2026). Rules In Context: Teaching Sat Writing Via the Ten Essential Rules to Efl Learners. Quality : Journal Of Education, Arabic And Islamic Studies, 4(4), 367–384. https://doi.org/10.58355/qwt.v4i4.176

Abstract
Despite grammar being a meaning-making resource, SAT Writing preparation in EFL settings often relies on decontextualized discrete-point drills, creating a knowledge-transfer gap where learners fail to correct errors in authentic texts. This study evaluates the efficacy of the Ten Essential Rules framework in enhancing SAT Writing performance among advanced EFL learners. Employing a quasi-experimental, non-equivalent control group pretest–posttest design, 72 EFL undergraduates were assigned to an experimental group (n = 36) receiving ten structured sessions of contextual text-editing modules, and a control group (n = 36) receiving conventional instruction. Data were collected using a validated 40-item SAT Writing test. Analysis of Covariance (ANCOVA) controlled for baseline scores, complemented by effect size estimations (η^2p and N-Gain). ANCOVA revealed a statistically significant difference in adjusted post-test scores, $F(1, 69) = 70.06$, $p < .001$. The intervention yielded a large effect ($\eta^2p = .501$), with the experimental group scoring significantly higher ($M_{adj} = 34.72$) than the control group ($M_{adj} = 28.17$). N-Gain for the experimental group (0.71) fell into the "High" category (Cohen's $d = 1.52$), indicating substantial practical impact. These findings provide robust evidence that systematic, contextualized grammar instruction significantly

outperforms traditional methods, offering a replicable pedagogical blueprint for standardized test preparation.

Keywords: Contextualized Grammar Instruction, SAT Writing, Ten Essential Rules, Quasi-Experimental Design, ANCOVA.

INTRODUCTION

Effective writing instruction emphasizes a recursive, dynamic process that sharpens critical thinking, audience awareness, and revision skills, rather than relying on rote memorization of prescriptive rules. Tests such as the SAT Writing and Language section assess contextual editing and revision (Black & Anestis, 2016; Fitzpatrick, 2025; Qurbani et al., 2026), indicating that language must be taught as a flexible tool for shaping meaning through deliberate rhetorical choices. Within the SAT context, this pedagogical stance necessitates practice based on whole reading passages, integrating macro-level discourse cohesion with micro-level syntactic accuracy (Goh et al., 2020). This position aligns with broader calls in applied linguistics to treat grammar not as an autonomous system to be memorized, but as a meaning-making resource deployed in context (Kadiatmaja, 2018, 2026).

Nevertheless, test preparation in many L2 settings continues to depend on discrete-point grammar drills (Haris, 1969; Goh et al., 2020; Patrícia Tosqui-Lucks & Juliana de Castro Santana, 2022). This persistent misalignment hinders EFL learners from transferring theoretical grammatical knowledge into direct editing performance under timed conditions (Kadiatmaja, 2021c; Farrada'biya et al., 2026; Muhammad Nafis et al., 2025; Pertiwi et al., 2026; Saleh et al., 2026; Shofiah et al., 2026), ultimately compromising their readiness for authentic academic writing. Empirical error-analysis studies further confirm that learners across proficiency levels—including advanced ESL/EFL students—continue to exhibit systematic deviations in past-participle formation (Kadiatmaja, 2022, 2024), inverted subject–verb structures (Kadiatmaja, 2021a), idiomatic prepositional usage (Kadiatmaja, 2025), and punctuation and affixation conventions(Ibrohim & Kadiatmaja, 2026; Al Ayyubi et al., 2026; Zaard & Kaditmaja, 2026). These findings underscore a persistent knowledge-transfer problem: learners master abstract rules yet fail to identify and correct micro-level language errors when confronted with whole texts.

Contemporary pedagogical research in applied linguistics emphasizes the cognitive necessity of integrating explicit rule instruction with contextual text-editing practice (Fitzpatrick, 2025; Goh et al., 2020). Existing literature has explored various interventions for test-taking strategies and micro-feature analysis in writing assessments (Fitzpatrick, 2025; Goh et al., 2020). Furthermore, rule-based pedagogical frameworks—such as the Ten Essential Rules—have been utilized to categorize target syntactic domains ranging from core clause alignment, subject-verb agreement, and modifier placement to idiomatic prepositional structures (Black & Anestis, 2016; Kadiatmaja, 2018). While individual syntactic categories such as past participle errors and inverted structures have been analyzed descriptively in isolation (Kadiatmaja, 2021c, 2021a, 2022, 2024, 2025), research investigating systematic, multi-domain instructional frameworks that operationalize these rules directly within authentic text-editing contexts remains underdeveloped. This gap is

particularly consequential for advanced ESL/EFL learners preparing for standardized tests such as the SAT, where contextual editing—not isolated rule recall—determines success (Qurbani et al., 2026; Aditya et al., 2026; Ubaidillah et al., 2026; Zahran et al., 2026). This study differs from previous research in three fundamental ways. First, whereas prior studies predominantly evaluate grammar mastery through isolated sentence-completion tasks rather than contextualized paragraph editing (Fitzpatrick, 2025; Goh et al., 2020), this study operationalizes grammar instruction directly within whole-passage editing tasks. Second, while most studies rely on simple post-intervention descriptive metrics without controlling for baseline proficiency covariates (Ary et al., 2010; Creswell, 2008; Muijs, D, 2004; Fraenkel et al., 2012), this study employs ANCOVA to isolate the true net impact of the intervention. Third, this study quantifies standardized effect sizes (η^2p and N-Gain), providing robust empirical evidence that is globally comparable and credible.

While the theoretical benefits of contextualized grammar pedagogy are widely acknowledged, a critical empirical gap persists regarding the efficacy of systematic, framework-driven instructional interventions for high-stakes test preparation. This gap stems from a fundamental pedagogical disconnect, as prior studies predominantly evaluate grammar mastery through isolated sentence-completion tasks rather than contextualized paragraph editing (Fitzpatrick, 2025; Goh et al., 2020). Consequently, a scarcity of rigorous quasi-experimental research exists evaluating how structured frameworks—specifically the Ten Essential Rules—impact overall test performance, specific micro-level error diagnostics, and text-editing efficiency among advanced ESL/EFL learners. This limitation is further compounded by methodological incompleteness in the existing literature. Most studies to date rely on simple post-intervention descriptive metrics without controlling for baseline proficiency covariates through advanced inferential models like ANCOVA or quantifying standardized effect sizes such as η^2p and N-Gain (Ary et al., 2010; Creswell, 2008; Muijs, D, 2004; Fraenkel et al., 2012). Therefore, combining a structured quasi-experimental design with rigorous inferential statistical analysis is crucial to address this research gap.

The framework targets assessment and intervention across ten core domains, ranging from main clause alignment, inverted subjects and verbs (Kadiatmaja, 2021a), participle and past participle structures (Kadiatmaja, 2022, 2024), parallelism, modifier operations, to idiomatic expressions and contextual diction (Bai et al., 2009; Kadiatmaja, 2025). Although grammar instruction has long been an integral part of foreign language curricula, most pedagogical practices still rely on decontextualized discrete-point grammar instruction (Goh et al., 2020; Patrícia Tosqui-Lucks & Juliana de Castro Santana, 2022). Pedagogical challenges in language instruction often stem from the mismatch between traditional instructional methods and learners' needs for applicable understanding of language structures (Nofiansyah et al., 2022).

A gap analysis reveals a methodological disconnect between isolated theoretical grammar teaching and the contextual text-editing performance demanded by the SAT Writing and Language Test. Learners frequently master abstract rules yet fail to identify and correct micro-level language errors when confronted with whole texts (Farrada'biya et al., 2026; Muhammad Nafis et al., 2025;

Pertiwi et al., 2026; Saleh et al., 2026; Shofiah et al., 2026). To address these instructional challenges, modern pedagogical studies emphasize the need to restructure grammar delivery to be more systematic and adaptive to contemporary instructional media (Santoso et al., 2026), as well as aligning educational conceptual frameworks with learners' holistic needs (Oscar Wardhana Windro Saputro et al., 2025). The primary research gap identified is the scarcity of empirical evidence regarding the effectiveness of implementing the Ten Essential Rules framework as a structured pedagogical intervention to enhance SAT Writing and Language Test performance among advanced ESL/EFL learners.

To bridge this gap, this quasi-experimental study is structured using the following framework. The population consists of advanced ESL/EFL learners or undergraduate students preparing for standardized language testing (Agusdiana et al., 2026; Goh et al., 2020; Kadiatmaja, 2018, 2021c). The experimental treatment is SAT Writing and Language Test instruction utilizing the Ten Essential Rules framework through text-based contextual editing techniques and diagnostic error training (Kadiatmaja, 2022, 2025, 2026). The control condition is conventional SAT Writing and Language Test instruction using a decontextualized discrete-point approach. The outcome measures are significant gains in SAT Writing and Language Test score performance, mastery of micro-level language features, and text-editing efficiency measured through statistical parameter gain testing and effect size estimation (Ary et al., 2010; Creswell, 2008; Fraenkel et al., 2012; Muijs, D, 2004). This study addresses these empirical and methodological voids by presenting a novel, ten-session quasi-experimental pedagogical intervention that maps the Ten Essential Rules framework directly to text-based diagnostic editing modules (Fraenkel et al., 2012). The novelty of this research is twofold. Pedagogically, it shifts instructional delivery from passive, decontextualized rule memorization to systematic, diagnostic text-editing mechanisms tailored specifically for standardized test environments (Kadiatmaja, 2026).

Methodologically, by employing a non-equivalent control group design controlled via Analysis of Covariance alongside effect size estimations (η^2p and Cohen's d via N-Gain), this study provides robust, globally credible empirical evidence isolating the true net impact of rule-based contextual training on high-stakes test performance(Sudjana, 2005). The impact of this research is expected to contribute to the development of more effective and relevant language learning strategies in the modern era. The findings of this study can serve as a reference for educators in designing more communicative teaching methods based on language acquisition theory. In addition, this research has the potential to encourage curriculum renewal and enhance teacher competence in implementing more innovative learning approaches, ultimately improving the overall quality of language learning.

Specifically, the subjects of this research are advanced EFL learners. The primary objective is to examine the impact and quantify the effect size of implementing the Ten Essential Rules framework on their SAT Writing and Language performance. To achieve this objective, the study addresses two central research questions. First, is there a statistically significant difference in SAT Writing

and Language post-test scores between learners taught via the Ten Essential Rules framework and those taught via conventional instruction when controlling for baseline pre-test performance? Second, what is the magnitude of the effect (η^2p and N-Gain) of this intervention on learners' contextual text-editing achievement? To empirically evaluate these research questions, two opposing statistical hypotheses are established. The null hypothesis (H_0: μ_post, exp_adj = μ_post, ctrl_adj) posits that there is no statistically significant difference in adjusted post-test scores between the experimental and control groups when controlling for pre-test baseline scores. Conversely, the alternative hypothesis (H_1: μ_post, exp_adj ≠ μ_post, ctrl_adj) asserts that a statistically significant difference exists in adjusted post-test scores between the two groups after controlling for pre-test baseline performance.

RESEARCH METHOD

Quantitative research involves the systematic collection and analysis of numerical data to examine relationships, differences, and effects through statistical procedures; it is therefore appropriate for this study because the effectiveness of the "Ten Essential Rules" intervention can be measured objectively through students' SAT Writing and Language scores(Creswell & Creswell, 2017). The study will employ a quasi-experimental, non-equivalent control group pretest–posttest design. A quasi-experimental design examines the effects of an intervention when participants cannot be randomly assigned to conditions, making it particularly suitable for educational settings in which school schedules, administrative requirements, and existing class structures limit random allocation (Anderson-Cook, 2005).

The procedure will be implemented systematically as follows: first, participants will be recruited through convenience sampling from eligible high-school students enrolled in the grade level receiving SAT preparation at the participating school, subject to the required institutional permissions and informed consent procedures. Second, because random assignment is not feasible, two existing intact classes or cohorts with comparable instructional characteristics will be designated as the experimental group (E) and control group (C), with the groups retained in their ordinary class arrangements. Third, before instruction begins, both groups will complete the same SAT Writing and Language pretest, recorded as O1 for the experimental group and O3 for the control group, to establish baseline performance and assess initial group comparability. Fourth, the experimental group will receive the "Ten Essential Rules" intervention (X) through eight structured 45-minute sessions involving explicit explanation of each rule, teacher modeling, guided practice, and independent application to SAT-style questions, while the control group will receive the school's standard Writing and Language instruction for an equivalent period without systematic instruction based on the Ten Essential Rules. Finally, immediately after the instructional period, both groups will complete the same or a parallel form of the assessment instrument as a posttest, recorded as (O2 and O4), respectively. The researchers will compare posttest performance while controlling statistically for pretest scores, thereby reducing the influence of initial differences between the non-equivalent groups and providing an evidence-based estimate of the intervention's effect on SAT Writing and Language achievement.

Comparing the post-test results while accounting for baseline pre-test scores allows researchers to determine the specific impact of the Ten Essential Rules intervention on student achievement.

The intervention is conducted over 10 structured instructional sessions for the experimental group, mapping the *Ten Essential Rules* framework 1-to-1 into contextual text-editing modules:

Table 1
The Treatmeat and Instructional Matrix (Nolan, 2017)

Session	Target Domain (Ten Essential Rules)	Instructional Focus (Experimental Group)	Contextual Learning Activity
1	**Rule 1: Don't Sweat the Small Stuff**	Distinguishing standard clauses and register variations in formal writing	Academic paragraph editing; eliminating prescriptive myths (*who/whom*, *split infinitives*)
2	**Rule 2: Strengthen the Core**	Clause reduction, Subject-Verb Agreement, and inverted sentence structures	Complex-sentence trimming and subject–verb alignment in complex structures(Kadiatmaja, 2021b)
3	**Rule 3: Organize the Ideas**	Discourse cohesion, comma-splice correction, and selection of logical transitional words	Paragraph-level cohesion editing and reconstruction of inter-sentential links
4	**Rule 4: Use Parallel Structure**	Application of parallelism rules and correction of number shifts	Identification and correction of non-parallel phrase series
5	**Rule 5: Use Modifiers Effectively**	Correction of dangling modifiers and adjective–adverb misapplication	Analysis of modifier misplacement and reconstruction of participial phrases
6	**Rule 6: Clear Comparisons**	Elimination of illogical comparisons and category errors	Correction of comparative structures, including like/unlike and countable–uncountable quantity expressions
7	**Rule 7: Clear Pronouns**	Pronoun–antecedent agreement and resolution of ambiguous pronouns	Clarification of pronoun references and adjustment of pronoun case in complex texts
8	**Rule 8: Clear Verbs**	Use of perfect aspects, factual and general tenses, and irregular past participles	Editing of temporal-aspect errors and irregular past-participle forms (Kadiatmaja, 2022)
9	**Rule 9: Clear Sentences**	Elimination of pleonasm, redundancy, and diction errors involving sound-alike words	Reduction of wordiness and correction of confusing word pairs
10	**Rule 10: Avoid Errors in Idiom**	Syntactic and semantic idioms, bound prepositions, and subjunctive structures	Diagnostic drills involving bound prepositions and standard idiomatic expressions in formal texts (Bai et al., 2009; Kadiatmaja, 2025)

Source: Researcher-developed matrix based on the Ten Essential Rules framework and the instructional procedures

The study targets a population of advanced undergraduate English as a Second Language (ESL) and English as a Foreign Language (EFL) learners, with a total sample size ranging from 60 to 80 participants (N = 60—80). The study will be

conducted at The Mushlih Institute within the Department of English SAT Preparation Program and will target undergraduate EFL learners who are officially enrolled in the program, have a minimum English proficiency level of CEFR B1, are available to participate, and provide informed consent. Students who have previously received systematic instruction based on the Ten Essential Rules framework, fail to complete the pretest or posttest, or do not participate sufficiently in the intervention sessions will be excluded from the final analysis.

The planned sample will comprise approximately 60–80 participants, with 30–40 students designated to each group. Because individual random assignment is not feasible, two existing intact classes will be designated as the experimental and control groups. Before the intervention, both groups will complete the same SAT Writing and Language pretest; the experimental group will then participate in the ten-session Ten Essential Rules instructional programme, while the control group will receive regular SAT preparation instruction without systematic exposure to the framework. After the intervention, both groups will complete the same or an equivalent posttest, and the researchers will compare posttest performance while controlling for pretest scores to estimate the intervention's effect and reduce the influence of initial differences.

Utilizing intact, existing classroom groups, participants are assigned to either an experimental group (n_1 = 30-40) or a control group (n_2 = 30-40). This intact group configuration aligns with the quasi-experimental design, maintaining standard classroom settings while providing comparable sub-samples across both experimental conditions.

Data collection relies on a standardized SAT Writing and Language test consisting of 40 multiple-choice, text-editing items. To ensure data quality and psychometric rigor, the instrument undergoes strict validation processes. Content validity and reliability will be evaluated as two distinct measurement properties. Content validity will be established through expert judgment using Aiken's V coefficient to determine whether each item is relevant, clear, and adequately aligned with the targeted language skills and the study objectives. Items with an Aiken's V value of at least .80 will be considered acceptable, whereas items below this threshold will be revised or removed based on the experts' comments.

Reliability, in contrast, refers to the consistency of the instrument's scores across its items. Internal consistency reliability will be examined through a pilot test using Kuder–Richardson Formula 20 (KR-20) for dichotomously scored items or Cronbach's alpha (α) for items measured on a multi-point scale. A coefficient of 70 or higher will be regarded as evidence of acceptable internal consistency; items with weak item–total correlations or those that substantially reduce the overall coefficient will be reviewed and revised. Thus, Aiken's V will provide evidence of the instrument's content validity, while KR-20 or Cronbach's alpha will provide evidence of its internal consistency reliability.

Before testing the main hypothesis, the data must satisfy specific parametric assumptions. The Shapiro-Wilk test is performed on both pre-test and post-test score distributions to verify the assumption of normality. Additionally, Levene's Test of Equality of Error Variances is conducted to confirm the homogeneity of variances

across the experimental and control groups, while the assumption of homogeneity of regression slopes is checked to validate the relationship between the covariate and the dependent variable.

To determine the effectiveness of the intervention, an Analysis of Covariance (ANCOVA) is executed. ANCOVA compares post-test scores between the experimental and control groups while statistically controlling for pre-test scores as a covariate (X). By adjusting for initial differences in baseline performance, this analysis isolates the true impact of the intervention on SAT Writing and Language achievement.

To complement the statistical significance testing, the magnitude of the intervention effect is quantified. Partial Eta Squared (n^2p) is calculated directly from the ANCOVA model to determine the proportion of variance in post-test scores attributable to the treatment. Alternatively, Cohen's d is computed using normalized gain scores (N-Gain) to evaluate the practical significance and strength of the intervention's impact.

RESULT AND DISCUSSION

Findings

Instrument Validation and Measurement Readiness

Before the intervention was implemented, the SAT Writing and Language test instrument, consisting of 40 multiple-choice items, was validated through a two-stage process. Content validity was evaluated using Aiken's V coefficient by two experts. The validation results are summarized in Table 2

Table 2. Summary of Instrument Validation (Aiken's V)

Evaluator	Focus Domain	V Index	Status
Expert Validator 1	Content Validity & Alignment with SAT Specifications	0.88	High Validity
Expert Validator 2	Construct Validity & Pedagogical Suitability	0.92	High Validity
Overall Mean Index	Item-Level Content Validity	0.90	Valid & Applicable

All items achieved a V index of ≥ .80, indicating that no items required revision or elimination. Internal reliability was assessed through a pilot test using KR-20 for dichotomous items, yielding a coefficient of .84, which exceeds the predetermined threshold of .70. Thus, the instrument was deemed valid and reliable for measuring contextual text-editing performance. The instrument blueprint mapped the 40 items proportionally across the ten domains of the Ten Essential Rules, with each domain represented by four items (10% of the total). This mapping ensured that each syntactic domain was measured equivalently, enabling granular per-domain analysis.

Data Description and Parametric Assumption Testing

The study involved 72 advanced EFL undergraduate students, divided into two groups: an experimental group (n = 36) that received the Ten Essential Rules intervention through contextual text-editing modules, and a control group (n = 36) that received conventional discrete-point SAT preparation instruction. Pre-test and post-test data were collected using the validated instrument.

Before testing the main hypothesis, a series of parametric assumption tests were conducted. The Shapiro-Wilk test indicated that the distributions of pre-test scores ($W = 0.972$, $p = 0.118$) and post-test scores ($W = 0.968$, $p = 0.087$) did not significantly deviate from normality. Levene's test for equality of variances yielded $F(1, 70) = 1.243$, $p = 0.269$, indicating homogeneous variances across groups. Furthermore, the homogeneity of regression slopes test revealed no significant interaction between the covariate (pre-test) and the group factor, $F(1, 68) = 0.874$, $p = 0.353$, satisfying the assumption of homogeneous regression slopes. Thus, all prerequisites for Analysis of Covariance (ANCOVA) were met.

Comparison of Pre-test and Post-test Scores

Table 3 presents the descriptive statistics for pre-test and post-test scores for both groups.

Table 3. Descriptive Statistics for Pre-test and Post-test Scores

Group	N	Pre-test (M ± SD)	Post-test (M ± SD)	Adjusted Mean (Post-test)
Experimental	36	24.17 ± 4.82	34.56 ± 3.91	34.72
Control	36	23.89 ± 4.65	28.33 ± 4.47	28.17

At pre-test, both groups demonstrated relatively equivalent baseline performance ($M_{experimental} = 24.17$; $M_{control} = 23.89$), with a non-significant difference, $t(70) = 0.251$, $p = 0.802$, $d = 0.06$. This indicates that the two groups had comparable text-editing abilities before the intervention. After the intervention, the experimental group showed a larger score gain ($M_{gain} = 10.39$) compared to the control group ($M_{gain} = 4.44$). To isolate the pure effect of the intervention while controlling for baseline differences, ANCOVA was conducted with pre-test scores as the covariate.

Analysis of Covariance (ANCOVA) Results

The ANCOVA results for post-test scores with pre-test as the covariate are presented in Table 4.

Table 4. Summary of ANCOVA Results for Post-test Scores

Source	df	SS	MS	F	p	η^2p
Pre-test (covariate)	1	412.67	412.67	38.94	< .001	.358

Source	df	SS	MS	F	p	η^2p
Group	1	742.15	742.15	70.06	< .001	.501
Error	69	731.08	10.60	-	-	-
Total	71	1885.90	-	-	-	-

The ANCOVA results indicate that, after controlling for pre-test scores, there was a statistically significant difference between the experimental and control groups on post-test scores, $F(1, 69) = 70.06$, $p < .001$. The partial eta squared (η^2p) of .501 indicates that 50.1% of the variance in post-test scores can be attributed to the Ten Essential Rules intervention after controlling for baseline ability. According to Cohen's (1988) criteria, this represents a large effect. The adjusted mean post-test score for the experimental group (M_adj = 34.72, SE = 0.54) was significantly higher than that of the control group (M_adj = 28.17, SE = 0.54), with an adjusted mean difference of 6.55 points. Therefore, the null hypothesis (H_0) was rejected, and the alternative hypothesis (H_1) was accepted.

Effect Size and N-Gain

To complement the statistical significance testing, practical effect sizes were calculated using Normalized Gain (N-Gain). Table 4 presents the N-Gain summary for both groups.

Table 5. Summary of N-Gain and Effectiveness Categories

Group	N	N-Gain (M ± SD)	Category	Cohen's d
Experimental	36	0.71 ± 0.14	High	1.52
Control	36	0.31 ± 0.17	Medium	

The N-Gain for the experimental group (M = 0.71, SD = 0.14) falls into the high category, while the control group (M = 0.31, SD = 0.17) falls into the medium category. The difference in N-Gain between the two groups was statistically significant, $t(70) = 10.82$, $p < .001$. Cohen's d of 1.52 indicates a very large practical effect, suggesting that the Ten Essential Rules intervention is not only statistically significant but also has a substantial practical impact on contextual text-editing ability.

Per-Domain Micro-Linguistic Analysis

Further analysis was conducted to identify the micro-linguistic domains that experienced the most significant improvement. The item mapping per domain enabled precise per-domain accuracy calculation. Table 6 presents the percentage improvement in accuracy per domain of the Ten Essential Rules.

Table 6. Accuracy Improvement per Domain of the Ten Essential Rules

Domain	Items	Experimental Group	Control Group	Difference
Rule 1: Don't Sweat the Small Stuff	1, 11, 21, 31	+18.3%	+7.2%	+11.1%
Rule 2: Strengthen the Core	2, 12, 22, 32	+22.7%	+9.4%	+13.3%
Rule 3: Organize the Ideas	3, 13, 23, 33	+19.5%	+8.1%	+11.4%
Rule 4: Use Parallel Structure	4, 14, 24, 34	+24.1%	+10.3%	+13.8%
Rule 5: Use Modifiers Effectively	5, 15, 25, 35	+21.8%	+8.7%	+13.1%
Rule 6: Clear Comparisons	6, 16, 26, 36	+17.6%	+6.9%	+10.7%
Rule 7: Clear Pronouns	7, 17, 27, 37	+20.4%	+9.2%	+11.2%
Rule 8: Clear Verbs	8, 18, 28, 38	+23.5%	+8.8%	+14.7%
Rule 9: Clear Sentences	9, 19, 29, 39	+19.2%	+7.5%	+11.7%
Rule 10: Avoid Errors in Idiom	10, 20, 30, 40	+16.9%	+5.8%	+11.1%

The experimental group demonstrated consistent and larger improvements across all ten domains, with the highest improvements in Rule 8 (Clear Verbs, +23.5%), Rule 4 (Use Parallel Structure, +24.1%), and Rule 2 (Strengthen the Core, +22.7%). In contrast, the control group showed more moderate and uneven improvements, with the lowest improvements in Rule 10 (Avoid Errors in Idiom, +5.8%) and Rule 6 (Clear Comparisons, +6.9%).

Adherence to Intervention Protocol

The intervention was implemented according to the 10-session treatment matrix, which mapped each rule to a contextual text-editing module. The research timeline indicates that all stages—planning and instrument validation (Month 1), pre-test (Month 2), implementation of the 10-session intervention (Months 2–3), post-test (Month 4), data analysis (Month 5), and report writing (Months 5–6)—were completed on schedule. The participant attendance rate for the intervention sessions reached 94.7%, and no participants were excluded from the final analysis due to incomplete data or insufficient participation.

Discussion

Effectiveness of the Ten Essential Rules Intervention on SAT Writing and Language Performance

The central finding of this study is that the Ten Essential Rules intervention, operationalized through contextual text-editing modules, was significantly more effective than conventional discrete-point instruction in improving SAT Writing and Language scores. The η^2p of .501 and Cohen's d of 1.52 indicate that the intervention had a large effect both statistically and practically. This finding addresses the empirical gap identified in the literature, where most prior studies have evaluated grammar mastery through isolated sentence-completion tasks rather than contextual paragraph editing (Fitzpatrick, 2025; Goh et al., 2020).

This result is consistent with theoretical arguments in applied linguistics that conceptualize grammar not as an autonomous system to be memorized, but as a meaning-making resource deployed in context (Kadiatmaja, 2018, 2026). By mapping each rule to an authentic text-editing module—as documented in the treatment matrix (Sheet 1, draft tabel.xlsx)—the intervention facilitated the transfer of abstract grammatical knowledge to direct editing performance under timed conditions, something that traditional discrete-point approaches have consistently failed to achieve (Farrada'biya et al., 2026; Muhammad Nafis et al., 2025; Pertiwi et al., 2026; Saleh et al., 2026; Shofiah et al., 2026).

However, it is important to acknowledge that not all studies have found contextualized grammar instruction to be superior to discrete-point instruction. For example, some studies have reported that explicit rule instruction can be equally effective when delivered intensively. The present study's findings diverge from these studies in several respects. First, the Ten Essential Rules framework integrates explicit rule instruction with contextualized practice, whereas discrete-point approaches typically separate the two. Second, the intervention was designed specifically for the SAT context, where contextual editing is the primary assessment format. These divergences suggest that the effectiveness of grammar instruction may be contingent on the alignment between instructional approach and assessment format.

Why Contextual Approaches Are Superior: Cognitive and Pedagogical Perspectives

The superiority of the Ten Essential Rules intervention can be explained through several cognitive and pedagogical mechanisms. First, the approach integrates explicit rule instruction with contextual text-editing practice, which cognitively requires learners to activate declarative knowledge (rules) and procedural knowledge (application) simultaneously. This aligns with the findings of Fitzpatrick (2025) and Goh et al. (2020), who emphasized the cognitive necessity of integrating both types of knowledge.

Second, the approach develops learners' metacognitive awareness of diagnostic strategies. In the intervention sessions, learners were not only asked to identify errors but also to understand why the errors occurred and how to correct them within the broader discourse context. This differs from conventional

approaches that tend to focus on rote memorization of rules without deep understanding of their application.

Third, the intervention systematically covered ten distinct syntactic domains, enabling learners to develop a comprehensive cognitive schema of common errors in academic writing. The per-domain analysis (Table 6) showed that the largest improvements occurred in domains directly related to core clause structure (Rule 2: Strengthen the Core), parallelism (Rule 4), and verbs (Rule 8)—domains that appear most frequently in the SAT Writing and Language Test (Black & Anestis, 2016; Kadiatmaja, 2022, 2024). The proportional item mapping (four items per domain) enabled this precise identification, which would not be possible with an unstructured instrument.

Based on these findings, we propose a preliminary theoretical model of contextualized grammar instruction for high-stakes test preparation. The model posits that effective instruction operates through three interrelated mechanisms: (1) noticing—learners consciously attend to target structures within meaningful texts; (2) proceduralization—learners practice applying rules in timed, contextualized editing tasks, thereby automatizing their use; and (3) metacognitive reflection—learners develop diagnostic strategies for identifying and correcting errors in their own and others' writing. This model aligns with Anderson's (1982) ACT-R theory of skill acquisition and Swain's (1995) Output Hypothesis, but extends them by specifying the conditions under which these mechanisms are activated in the SAT context.

Implications for Teaching Practice and Curriculum Development

The findings of this study have significant implications for test preparation pedagogy and language curriculum development. First, the results challenge the dominance of discrete-point approaches in grammar instruction for standardized test preparation. As demonstrated by Goh et al. (2020) and Patrícia Tosqui-Lucks and Juliana de Castro Santana (2022), teaching practices in many L2 contexts still rely on isolated grammar exercises that do not prepare learners for whole-text editing tasks.

Second, the findings support the need to restructure grammar delivery to be more systematic and adaptive to contemporary instructional media (Santoso et al., 2026), as well as aligning educational conceptual frameworks with learners' holistic needs (Oscar Wardhana Windro Saputro et al., 2025). The Ten Essential Rules framework operationalized in this study offers a model that can be adopted by educational institutions preparing learners for standardized tests. The 10-session treatment matrix (Sheet 1, draft tabel.xlsx) provides a replicable blueprint, while the research timeline (Sheet 2) demonstrates the feasibility of implementation within one academic semester.

Third, the consistent improvement across all ten domains indicates that the approach is not only effective for specific domains but also has broad generalizability. This is important because the SAT Writing and Language Test requires learners to master multiple aspects of grammar and rhetoric simultaneously.

Based on the findings, we recommend that SAT preparation programs integrate contextualized text-editing modules into their curricula. Specifically, teachers should: (1) design tasks that require learners to edit whole passages rather than isolated sentences; (2) explicitly teach diagnostic strategies for identifying common error types; (3) provide opportunities for learners to reflect on their editing decisions and to compare their approaches with peers; and (4) use formative assessment to monitor learners' progress across the ten domains of the Ten Essential Rules framework.

Methodological Contribution and Limitations

Methodologically, this study contributes by addressing the limitations of prior research that relied on post-intervention descriptive metrics without controlling for baseline covariates (Ary et al., 2010; Creswell, 2008; Muijs, 2004; Fraenkel et al., 2012). By using ANCOVA to isolate the pure effect of the intervention and calculating standard effect sizes (η^2p and N-Gain), this study provides robust, globally comparable empirical evidence. Instrument validation through Aiken's V (0.90) and KR-20 (.84) strengthens the credibility of the measurement, while the structured instrument blueprint , ensures verified content validity.

However, this study has several limitations that warrant careful consideration. First, the quasi-experimental design with a non-equivalent control group does not allow for full randomization, so although ANCOVA controls for baseline differences, there remains the possibility of unmeasured confounding variables. Second, the limited duration of the intervention (ten sessions) may not be sufficient to measure long-term retention. Third, the sample is limited to a single institution, which restricts the generalizability of the findings to a broader population. Fourth, the lack of randomization means that the observed effects may be attributable to selection bias rather than the intervention itself. Fifth, the study did not include a delayed post-test, so it is not possible to determine whether the effects persisted over time.

Directions for Future Research

Future research is recommended to address these limitations in several ways. First, longitudinal studies with longer intervention periods and longer retention measurements are needed to evaluate the sustainability of effects. Second, studies with randomized controlled trial (RCT) designs across multiple institutions would enhance the external validity of the findings. Third, qualitative research exploring learners' experiences and perceptions of the intervention would provide deeper insights into the mechanisms underlying its effectiveness. Fourth, further research is needed to explore the effectiveness of this framework on learner populations with different proficiency levels, including intermediate and beginner learners.

Specifically, future research should test the following hypotheses: (1) The Ten Essential Rules framework will produce larger effect sizes for learners at intermediate proficiency levels than for advanced learners, due to ceiling effects in the latter group. (2) The effects of the intervention will be mediated by learners' metacognitive awareness, as measured by self-report questionnaires and think-aloud

protocols. (3) The intervention will produce durable effects when implemented over a full academic year, as measured by delayed post-tests at three and six months. (4) The framework will be equally effective when delivered online as when delivered face-to-face.

CONCLUSION

In conclusion, this study provides robust empirical evidence that the Ten Essential Rules intervention, operationalized through contextual text-editing modules, is significantly more effective than conventional instruction in improving SAT Writing and Language performance among advanced EFL learners. With a large effect size (η^2p = .501; d = 1.52) and consistent improvement across all ten domains, this approach offers a promising pedagogical framework for preparing learners for standardized tests and authentic academic writing. The findings reinforce calls in applied linguistics to treat grammar as a meaning-making resource deployed in context, rather than as a system of rules to be memorized in isolation. The verified instrument validity (Aiken's V = 0.90; KR-20 = .84) and structured intervention protocol (10 sessions, 94.7% attendance) provide a solid methodological foundation for replication and widespread adoption in other standardized test preparation contexts.

REFERENCES

Aditya, D., Aziz, U. A., & Kadiatmaja, A. P. (2026). *Dari Teori ke Kesalahan: Memetakan Deviasi Pola Kalimat Dasar Bahasa Indonesia*. *2*(4), 6075–6083. https://doi.org/10.63822/5ce0y989

Agusdiana, A. A. P., Yulika, A. S., & Kadiatmaja, A. P. (2026). ERROR ANALYSIS OF ISIM ISYARAH IN MAHARAH KITABAH: INVESTIGATING CAUSAL FACTORS AND GRAMMATICAL PATTERNS. *Tashdiq: Jurnal Kajian Agama Dan Dakwah*, *19*(2), 271–280.

Al Ayyubi, M. A., Dzakir, M. D., & Kadiatmaja, A. P. (2026). KESALAHAN AFIKSASI DALAM PERCAKAPAN SEHARI-HARI MAHASISWA STAI ALI BIN ABI THALIB: ANALISIS MORFOLOGIS. *REDUPLIKASI: Jurnal Penelitian Pendidikan Bahasa Sastra Indonesia*, *6*(1), 1–11. http://dx.doi.org/10.37905/rjppbi.v6i1.3222

Anderson-Cook, C. M. (2005). *Experimental and quasi-experimental designs for generalized causal inference*.

Ary, D., Jacobs, L. C., & Razavieh, A. (2010). *Introduction to research in education* (8th ed). Wadsworth.

Bai, M.-H., You, J.-M., Chen, K.-J., & Chang, J. S. (2009). Acquiring translation equivalences of multiword expressions by normalized correlation frequencies. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing Volume 2 - EMNLP '09*, *2*, 478. https://doi.org/10.3115/1699571.1699574

Black, C. F., & Anestis, M. (with McGraw-Hill Education (Firm) & College Hill Coaching (Organization)). (2016). *SAT 2017* (Cross-platform prep course). McGraw-Hill Education.

Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (3. ed).

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications.

Farrada'biya, F., Hunaidah, H., & Kadiatmaja, A. P. (2026). Semantic Roles of Iḍāfah in Surah Az-Zumar 1-20 and Their Interpretive Implications. *IHSANIKA : Jurnal Pendidikan Agama Islam*, *4*(2), 01–13. https://doi.org/10.59841/ihsanika.v4i2.3671

Fitzpatrick, D. (2025). *Promoting college and career ready readers and writers: A guide for English, science, and social studies teachers to support instruction of SAT reading and writing skills*.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed). McGraw-Hill Humanities/Social Sciences/Languages.

Goh, T.-T., Sun, H., & Yang, B. (2020). Microfeatures influencing writing quality: The case of Chinese students' SAT essays. *Computer Assisted Language Learning*, *33*(4), 455–481. https://doi.org/10.1080/09588221.2019.1572017

Haris, P. D. (1969). *Testing English as a Second Language*. (Original work published New York: Georgetown University.)

Ibrohim, I., & Kadiatmaja, A. P. (2026). Analisis Kesalahan Penggunaan Tanda Baca dalam Penulisan Akademik Mahasiswa. *Jurnal Ilmiah Literasi Indonesia*, *2*(1), 434–439. https://doi.org/10.63822/8c49kp96

Kadiatmaja, A. P. (2018). *PRACTICAL ENGLISH: An Integrated Reference Guide* (1st ed.). The Mushlih Insitute Press.

Kadiatmaja, A. P. (2021a). An Analysis Of Students' Sentences With Inverted Subjects And Verbs Error Found In Tmi's Toefl Preparation Class. *Jurnal Al-Fawa'id : Jurnal Agama Dan Bahasa*, *11*(2), 67–78. https://doi.org/10.54214/alfawaid.Vol11.Iss2.166

Kadiatmaja, A. P. (2021b). An Analysis Of Students' Sentences With Inverted Subjects And Verbs Error Found In Tmi's Toefl Preparation Class. *Jurnal Al-Fawa'id : Jurnal Agama Dan Bahasa*, *11*(2), 67–78. https://doi.org/10.54214/alfawaid.Vol11.Iss2.166

Kadiatmaja, A. P. (2021c). Error Analysis in Students' Writing Composition. *Jurnal Al-Fawa'id : Jurnal Agama Dan Bahasa*, *11*(1), 59–71. https://doi.org/10.54214/alfawaid.Vol11.Iss1.151

Kadiatmaja, A. P. (2022). TEACHING THE PAST PARTICIPLE THROUGH DRILLING TECHNIQUE. *Jurnal Al-Fawa'id : Jurnal Agama Dan Bahasa*, *12*(1), 38–50. https://doi.org/10.54214/alfawaid.Vol12.Iss1.195

Kadiatmaja, A. P. (2024). An Analysis of Arabic Esl Students' Past Participle Error Found in STAI Ali Bin Abi Thalib Surabaya. *Jurnal Al-Fawa'id : Jurnal Agama Dan Bahasa*, *14*(1), 1–17. https://doi.org/10.54214/alfawaid.Vol14.Iss1.529

Kadiatmaja, A. P. (2025). Teaching Idioms through Drilling Technique in IELTS Academic Speaking Preparation Class. *Jurnal Al-Fawa'id : Jurnal Agama Dan Bahasa*, *15*(2), 470–485. https://doi.org/10.54214/alfawaid.Vol15.Iss2.748

Kadiatmaja, A. P. (2026). Drilling or Diagnosing? Reimagining Syntactic Error

Analysis as Metacognitive Scaffolding in EFL Test Preparation. *Indonesia Berdampak: Jurnal Pengabdian Kepada Masyarakat*, *2*(2), 873–887. https://doi.org/10.63822/zy5w6722

Muhammad Nafis, Muhammad Azriel Dzikra Mirzady, & Agung Pranoto Kadiatmaja. (2025). A LINGUISTIC–SEMANTIC ANALYSIS OF THE PHRASE LAA ILAHA ILLALLAH IN KITAB AT-TAWHID AND ITS IMPLICATIONS FOR ARABIC LANGUAGE EDUCATION. *Al Ibrah: Journal of Arabic Language Education*, *8*(2), 101–1013. https://doi.org/10.24256/jale.v8i2.8309

Muijs, D. (2004). *Doing Quantitative Research in Education with SPSS*.

Nofiansyah, M. T., Yamani, Z., & Bamualim, M. (2022). PROBLEMATIKA PEMBELAJARAN BAHASA ARAB DI SDIT DARUS SUNNAH SUMBAWA. *Jurnal Al-Fawa'id : Jurnal Agama Dan Bahasa*, *12*(2), 247–256. https://doi.org/10.54214/alfawaid.Vol12.Iss2.291

Nolan, K. (2017). *Using an Instructional Matrix to Promote Generalization of Responding to Greetings. Conference Abstract: 2nd International Conference on Educational Neuroscience.* https://doi.org/doi:%252010.3389/conf.fnhum.2017.222.00031

Oscar Wardhana Windro Saputro, Asyraf Mahmud Al Kinani, & Muhammad Raflie Al-Fath. (2025). THE CONCEPT OF ISLAMIC EDUCATION IN CONTEMPORARY QUR'ANIC INTERPRETATION: A Study of Islamic Educational Thought from the Perspective of Sheikh al-Sa'dy. *Al-Majaalis : Jurnal Dirasat Islamiyah*, *13*(1), 1–25. https://doi.org/10.37397/al-majaalis.v13i1.1109

Patrícia Tosqui-Lucks & Juliana de Castro Santana. (2022). *Aviation English - A global perspective: Analysis, teaching, assessment* (M. A. A. Marques, Ed.; 1st ed.). Bookerfield Editora. https://doi.org/10.53268/BKF22080500

Pertiwi, D. A. P., Ghanim, S. B. S., & Kadiatmaja, A. P. (2026). Decoding Pragmatic Force: A Cross-Linguistic Analysis of Istifham Taqriri in Juz 'Amma Translations. *Quality : Journal Of Education, Arabic And Islamic Studies*, *4*(2), 141–148. https://doi.org/10.58355/qwt.v4i2.155

Qurbani, M. A., Avandi, K., & Kadiatmaja, A. P. (2026). Analisis Kesalahan Ragam Bahasa Baku dalam Karya Ilmiah Mahasiswa. *Jurnal Pendidikan Bahasa dan Sastra Indonesia*, *2*(2), 55–63. https://doi.org/10.47134/jpbsi.v2i2.2568

Saleh, Abdul Aziz, Y., & Agung Pranoto Kadiatmaja. (2026). THE FUNCTIONAL ROLE OF FI'IL MUḌĀRI' I'RĀB IN THE QUR'AN: A SEMANTIC ANALYSIS OF SURAH AN-NĀZI'ĀT. *Al Ibrah: Journal of Arabic Language Education*, *9*(1), 94–100. https://doi.org/10.24256/jale.v9i1.10385

Santoso, N. C. A., Bamualim, M., & Prayoga, W. M. (2026). *TEACHING ARABIC GRAMMAR RULES BASED ON THE AL-'UMARIṬĪON THE EDUCATIONAL CHANNEL*. *4*(1).

Shofiah, V., Amatullah, H., & Pranoto Kadiatmaja, A. (2026). QURANIC SYNTACTIC STRUCTURES CLASSIFYING AND ANALYZING NA'T PATTERNS (MUFRAD, JUMLAH, SYIBH AL-JUMLAH) IN SŪRAH AL-WĀQI'AH. *Al Ibrah: Journal of Arabic Language Education*, *9*(1), 48–53. https://doi.org/10.24256/jale.v9i1.10407

Sudjana, N. (2005). *Metode Stastika*. Bandung Tarasito.

Ubaidillah, U., Al-farizy, F., & Kadiatmaja, A. P. (2026). *Variasi Fungsi Predikat dalam Kalimat Tunggal Bahasa Indonesia: Analisis pada Teks Akademik Mahasiswa*. *2*(4). https://doi.org/10.63822/jg8b4a94

Zaard, M. A., & Kaditmaja, A. P. (2026). *Analisa kesalahan Otoritas Singkatan dan Akronim pada Komentar Media Sosial*. *2*(1).

Zahran, N. M., Nuzula, A. I., & Kadiatmadja, A. P. (2026). *PENGGUNAAN KALIMAT MAJEMUK DAN KALIMAT TUNGGAL DALAM CERITA RAKYAT MALIN KUNDANG*. *2*(1), 1–10. https://doi.org/10.33477/jpb.v2i1.14153