ARCHAIA JOURNAL OF SCIENCE AND TECHNOLOGY
VOL. 2, NO. 3, Sep 2025
https://doi.org/10.48161/ajst.v2n3a91

The Influence of Digital Communication on Grammar, Spelling, and Formal Writing Skills

Irfan Nurhidayat[1*]

[1] Department of Primary Education, Graduate Program, Majalengka University, Majalengka, Indonesia;

* **Corresponding author:** irfannurhidayat09@gmail.com.

ABSTRACT: Digital communication has expanded the frequency and social range of everyday writing while normalizing abbreviations, compressed syntax, non-standard spelling, emojis, and unconventional punctuation. This study examined participant-level relationships between digital-writing practices and formal-writing outcomes and evaluated an eight-week register-shifting and proofreading intervention. We allocated 280 university students from three institutions equally to an intervention or equal-contact comparison condition; 254 participants completed the follow-up assessment. At baseline and follow-up, participants contributed privacy-screened digital-message samples, an academic essay, a formal email, questionnaire data, an Online Placement Test score mapped to CEFR bands, and a register-awareness task. Digital-feature density predicted a higher rate of comparable forms in formal writing (rate ratio = 1.28, 95% CI [1.11, 1.46]) but did not independently predict overall writing quality ($\beta = 0.05$, 95% CI [-0.03, 0.12], $p = .203$). Register awareness attenuated the association between digital density and comparable formal forms, whereas the proofreading interaction did not remain significant after multiplicity adjustment. Condition-by-time effects were 0.18 for grammar, 0.33 for spelling and mechanics, 0.48 for register appropriateness, and 0.19 for overall writing quality. Framework analysis of 36 interviews identified audience-sensitive shifting, automaticity under time pressure, revision as audience redesign, and negative cases in which weak writing occurred without textisms. The findings supported a register-mediated account: digital forms differed strongly by context, and the most practically meaningful instructional gain concerned register appropriateness rather than overall writing quality.

Keywords: digital communication; formal writing; register awareness; grammar and spelling; digital literacy.

I. INTRODUCTION

Digital communication has changed the scale, speed, and social distribution of everyday writing. University students move repeatedly among private messaging, group chats, social-network posts, learning-management systems, institutional email, and assessed academic prose. These environments differ in audience, purpose, permanence, time pressure, shared knowledge, interpersonal distance, and consequences for linguistic choice. Abbreviations, clipped forms, omitted sentence closure, repeated punctuation, emojis, phonetic spellings, and unconventional capitalization can therefore function as efficient, socially meaningful resources rather than accidental failures. Corpus-based research shows that non-standard forms are shaped by communicative purpose, recipient relationship, and sender characteristics [1], [2].

Grammar, spelling, punctuation, coherence, lexical precision, and register appropriateness influence how academic and professional texts are evaluated, and highly visible digital forms may be interpreted as evidence of weak language control. However, existing evidence does not support a simple deterioration account. Studies of texting and conventional grammar have reported mixed and feature-specific relationships [3]–[5]. Additional evidence indicates that the pattern can vary with language and literacy

context [6]. Cohort and perception research has also shown that textese remains a recognizable register and that readers' evaluations depend on audience expectations and feature density [7], [8]. The central issue, therefore, is not exposure alone, but whether writers recognize contextual demands and regulate their linguistic choices.

Digital participation can also support fluency, audience awareness, feedback, and genre flexibility. Research on second-language social writing and computer-mediated collaboration has identified benefits when activities are embedded in explicit pedagogy [9]–[11]. Reviews of digital support systems similarly emphasize the importance of instructional design [12]. Multimodal composing research emphasizes audience and semiotic choice [13]–[15], although the present study focused on plain-text messages and formal written tasks. The pedagogical rationale, therefore, was not that technology automatically improves formal writing, but that contrastive activities can make audience, genre, and revision decisions more visible to learners.

The empirical literature remains difficult to interpret because many studies rely on self-reported screen time, unrelated corpora, or a single holistic writing score. Such designs cannot determine whether a writer's digital practices align with their formal-language choices, and they may confound digital habits with proficiency, first language, discipline, reading exposure, autocorrection, and proofreading behavior. Unpaired corpora also make it impossible to distinguish population differences from register differences. A stronger design links authentic digital and formal texts at the participant level, repeats measurements over time, and separates comparable surface features from multidimensional writing quality.

A further weakness concerns intervention evidence. Digital-writing activities often lack an equal-contact comparison condition, documented allocation, fidelity monitoring, or blinded outcome rating. When improvement occurs, it may reflect additional writing time, instructor attention, novelty, or peer interaction rather than the intended instructional mechanism. Accordingly, register-focused instruction requires a design that separates the observational digital-to-formal association from the condition-by-time instructional effect and reports both statistical and practical significance.

The present study addressed these limitations through a three-site, paired longitudinal mixed-methods intervention. Each participant supplied a naturalistic digital-message sample and completed controlled academic and professional writing tasks at baseline and follow-up. We measured register awareness, proofreading, proficiency, and writing-support-tool use directly. The intervention emphasized audience recognition, expansion of compressed forms, sentence-boundary repair, tone, cohesion, and staged proofreading, whereas the comparison condition received equal contact time and general academic-writing practice.

The study's originality was bounded to three contributions. First, it offered a design contribution by linking digital and formal registers within the same writers. Second, it offered an analytical contribution by separating comparable digital-associated forms from global writing quality. Third, it offered an intervention contribution by estimating whether explicit register-shifting and proofreading instruction improved contextual control over time. This structure allowed the study to examine influence as a conditional, statistically estimated relationship rather than automatic causation.

II. RESEARCH PROBLEM

Conceptually, public and educational debate often treats abbreviations, omitted punctuation, shortened words, emojis, and lower-case forms as evidence of declining language standards. This interpretation neglects register: a form that is efficient and affiliative in peer messaging may be inappropriate in assessed academic prose without being intrinsically defective.

Measurement is also problematic because digital language and formal writing are multidimensional. Avoiding textisms does not guarantee grammatical accuracy, lexical development, cohesion, coherence, or effective task response. Conversely, an isolated informal form does not invalidate an otherwise successful text. Research therefore requires separate indicators of digital features, grammatical and mechanical accuracy, lexical resources, syntactic control, cohesion, development, and register appropriateness.

Design limitations remain decisive. Unpaired samples cannot establish within-writer relationships, cross-sectional studies cannot evaluate change, and intervention claims require documented allocation, fidelity,

attrition, and repeated outcomes. The study therefore paired registers, repeated measurement, used count-appropriate and mixed-effects models, and linked qualitative explanations to participant identifiers and transcript lines.

III. RESEARCH OBJECTIVES

The study examined participant-level relationships between digital-language features and formal-writing outcomes and evaluated whether explicit register-shifting and proofreading instruction improved formal accuracy and register appropriateness over time.

- To quantify register-sensitive features in participants' digital messages, academic essays, and formal email tasks.
- To distinguish the selective occurrence of comparable digital-associated forms in formal writing from multidimensional overall writing quality.
- To test register awareness and proofreading as prespecified moderators while treating age, first language, discipline, platform, and tool use as exploratory variables.
- To estimate condition-by-time effects on grammar, spelling and mechanics, register appropriateness, comparable-form occurrence, and overall writing quality.
- To integrate quantitative outcomes with line-referenced interview evidence concerning audience recognition, automaticity, digital identity, and revision decisions.

IV. RESEARCH QUESTIONS

- RQ1. How did the same participants' digital messages, academic essays, and formal emails differ in register-sensitive linguistic features?
- RQ2. To what extent did participant-level digital-feature density predict comparable digital-associated forms and overall quality in formal writing?
- RQ3. Did register awareness and proofreading moderate the association between digital-feature density and comparable forms in formal writing?
- RQ4. Did the register-shifting and proofreading condition produce greater baseline-to-follow-up improvement than the equal-contact comparison condition?
- RQ5. How did interviewed participants explain audience recognition, automatic habits, identity, autocorrection, and revision across digital and formal contexts?

V. LITERATURE REVIEW

1. DIGITAL COMMUNICATION AS A SOCIOLINGUISTIC REGISTER

Computer-mediated communication is shaped by speed, recipient familiarity, turn-taking, privacy, shared context, and platform affordances. Corpus research indicates that digital messages combine standard and non-standard resources and that feature use varies by communicative purpose [1], [2]. Longitudinal and perception studies further show that textese remains patterned by cohort and audience expectations [7], [8]. Register analyses distinguish situationally conditioned language patterns from genre and style [16]. Register theory is therefore more informative than a deficit approach because it asks what a form accomplishes in a recurring situation. Naturalistic message research also faces privacy-driven self-selection, making sampling instructions, coverage indicators, and exclusion logs central to validity.

2. GRAMMAR, SPELLING, AND PARTICIPANT LINKAGE

Studies differ in age group, language, platform, data source, and outcome definition. Actual-text paired designs can estimate whether comparable forms occur across registers within writers, whereas self-report studies primarily estimate associations with perceived use. Multilingual writers introduce additional complexity because first-language influence, educational history, and proficiency can resemble digital transfer. Evidence from school-writing and spelling studies remains mixed and feature-specific [3]–[5]. Related research in Spanish and university settings further suggests that language context and self-reported

use matter [6], [9]. These findings support feature-level analysis with direct proficiency and language-background controls rather than a single exposure score.

3. FORMAL WRITING AS A MULTIDIMENSIONAL CONSTRUCT

Formal-writing quality extends beyond correctness. Lexical sophistication, syntactic complexity, and discourse development contribute to performance [17]–[19]. Measures of accuracy and discourse quality also vary in construct coverage and methodological implementation [20]–[22]. Corpus-derived syntactic measures provide an additional perspective but remain sensitive to task and genre [23]. A participant may avoid digital-associated forms yet produce weak organization or limited argumentation; another may include one informal contraction while producing coherent and lexically precise prose. The study therefore separated comparable digital-form occurrence from overall analytic quality.

4. REGISTER-FOCUSED INSTRUCTION AND LONGITUDINAL EVIDENCE

Research on collaborative writing and digital support suggests that explicit attention to audience and revision may improve writing [11], [12], [24]. Work on genre knowledge and focused written corrective feedback further emphasizes metacognition and focused formative support [25]–[27]. Writing-process theory provides a complementary account of reviewing and revision [28]. This literature matters because it identifies mechanisms such as contrastive analysis, focused proofreading, audience modeling, visible revision, and explanations of formality decisions. Intervention evidence is strongest when the comparison condition is active, contact time is equivalent, fidelity is documented, and rating is blind to condition and wave.

5. CRITICAL SYNTHESIS

The literature supports three bounded conclusions. Digital-associated forms are patterned and context-sensitive; their relationships with conventional writing are generally small or mixed once proficiency and broader literacy resources are considered; and evidence is strongest when researchers analyze longitudinal data from the same writers and distinguish intervention effects from observational associations. These conclusions guided the participant-linked design, the selection of moderators, and the interpretation of practical significance.

Table 1. Critical comparison of evidence informing the study.

Study/strand	Sample and context	Design/source	Measures	Major finding	Limitation	Relevance
Lyddy et al. [2]	University students; Ireland; English	Naturalistic SMS analysis	Non-standard forms	Systematic audience-sensitive variation	Convenience and message-selection bias	Message sampling and privacy
van Dijk et al. [3]	School learners; Netherlands; Dutch	Texting and grammar study	Texting features; grammar	Conditional rather than uniformly adverse relationships	Different age group	Control prior linguistic ability
Verheijen et al. [4], [5]	Youth; Netherlands; Dutch	Social-media and school-writing studies	Feature transfer; spelling	Transfer varied by feature and task	Context-specific measures	Feature-level and task-sensitive analysis

Study/strand	Sample and context	Design/source	Measures	Major finding	Limitation	Relevance
Gómez-Camacho et al. [6]	Spanish writers; Spanish	Orthographic analysis	Textisms and spelling	Forms were structured and phonologically motivated	Language-specific system	Multilingual caution
Parrella et al. [9]	University students; USA; English	Correlational self-report study	Digital use; perceived ability	Association was nuanced	Perceived outcome	Actual-text measurement
Li [11]; Strobl et al. [12]	International L2 research	Reviews of collaborative and digital writing	Revision; feedback; development	Benefits depended on pedagogy and task design	Not direct tests of the present programme	Intervention rationale
Polio and Shea [20]; Graesser et al. [21]	Educational corpora; multiple contexts	Computational and corpus-based writing analysis	Lexical and developmental indices	Writing quality was multidimensional and task-sensitive	Genre and proficiency sensitivity	Separate quality dimensions
Negretti and McGrath [25]; Hayes [28]	Advanced writing and process research	Instructional and process studies	Genre knowledge; revision	Reflection and revision strengthened control	Different populations and tasks	Awareness and proofreading mechanisms
Polio and Shea [20]; Graesser et al. [21]	Educational writing corpora	Accuracy and discourse measurement	Accuracy; cohesion; discourse	Measures differed in construct coverage	Method-dependent estimates	Measurement validation
Crossley and McNamara [22]; Lu [23]	L2 writing corpora	Corpus-based syntactic analysis	Syntactic complexity	Corpus-derived indices added complementary evidence	Task and genre sensitivity	Sensitivity analysis

VI. RESEARCH GAP AND STUDY CONTRIBUTION

The study addressed a participant-linkage gap by pairing digital and formal texts, an analytical gap by separating comparable digital-form occurrence from overall quality, and an intervention gap by evaluating explicit register-shifting and proofreading instruction over time.

Its contribution was descriptive, analytical, and methodological. Descriptively, it documented the extent of register differentiation within the same writers. Analytically, it estimated whether digital-feature density was related to comparable formal forms after proficiency, register awareness, and proofreading were considered. Methodologically, it combined repeated quantitative measurement with traceable qualitative evidence and a prespecified intervention contrast.

VII. THEORETICAL FRAMEWORK AND ANALYTICAL MODEL

Register theory and writing-process theory formed the primary framework. Register theory explained why linguistic choices varied with audience, purpose, channel, and institutional expectations [16]. Writing-process theory located formal-language control within task interpretation, formulation, reviewing, and revision [28]. Usage-based learning and sociolinguistic style shifting served as supporting explanations:

frequent digital routines may remain accessible, while writers may select or inhibit them to construct stance and audience relationships.

Proficiency represented general linguistic capacity, register awareness represented contextual recognition, and proofreading represented revision behaviour. The study directly tested the relationships linking digital-feature density, formal-writing outcomes, register awareness, proofreading, proficiency, and instructional condition.

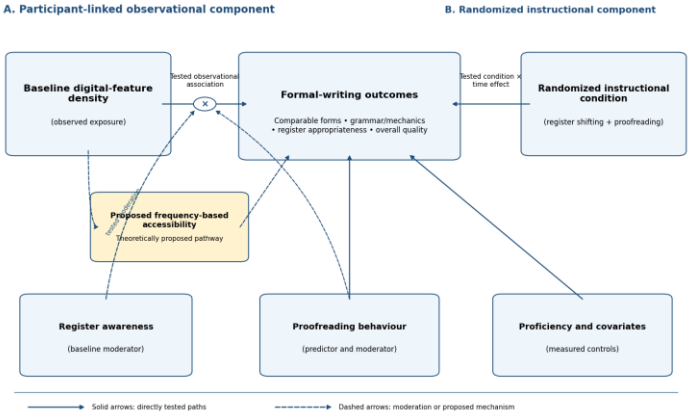


FIGURE 1. Register-mediated framework showing direct, moderation, and theory-based explanatory paths.

Table 2. Theoretical constructs, measures, and analytical roles.

Construct	Core meaning	Measure	Measurement status	Analytical role
Register theory	Situationally conditioned variation	Register-awareness task; appropriateness rating	Directly measured	Primary explanation of contextual selection
Writing-process theory	Task interpretation and reviewing	Proofreading scale; revision account	Directly measured and qualitatively examined	Revision behaviour
Usage-based learning	Frequency-based accessibility	Digital-feature density	Approximated	Supporting explanation
Sociolinguistic style shifting	Audience design and identity	Interview coding	Qualitatively examined	Supporting explanation
Proficiency	General linguistic capacity	OOPT score mapped to CEFR band	Directly measured	Covariate and quality predictor
Instructional condition	Explicit contrastive instruction	Condition, time, attendance, fidelity	Directly measured	Tests instructional change

VIII. METHODOLOGY

1. RESEARCH DESIGN

The study used a two-arm, three-site, paired longitudinal mixed-methods intervention design with baseline and follow-up measurement. The digital-to-formal relationship was observational at participant level, whereas the instructional effect was estimated from concealed 1:1 allocation, equal contact time, masked outcome rating, fidelity monitoring, and intention-to-treat analysis. Register appropriateness was

the primary instructional outcome; grammar, spelling and mechanics, comparable digital-associated forms, and overall writing quality were secondary outcomes. Framework analysis followed the matrix-based procedure in [29], with trustworthiness procedures informed by [30]. Linear mixed-effects estimation and associated inference followed established implementations [31], [32]. Effect-size reporting and conventional power checks followed published guidance [33], [34]. We interpreted inter-rater agreement using established kappa guidance [35], and we integrated mixed methods following joint-display principles [36]. Trial reporting followed CONSORT guidance [37], and negative-binomial sensitivity models used a flexible generalized mixed-model framework [38].

2. SETTING, PARTICIPANTS, SAMPLING, AND POWER

We conducted fieldwork at three public universities, reported as Sites A, B, and C, to preserve institutional confidentiality. Recruitment took place from 8 February to 12 March 2021 through course announcements, institutional email lists, and student information sessions. Eligible participants were at least 18 years old, enrolled in undergraduate or postgraduate study, used English in digital communication on at least three days per week, and could provide the required participant-authored text. Students were excluded when message coverage was insufficient, baseline tasks were incomplete, third-party privacy could not be protected, or prohibited writing-support tools were used during controlled tasks.

We invited 360 students, screened 322, found 294 met the eligibility criteria, and had 280 provide consent and complete the baseline assessment. Participants represented humanities and social sciences, business, and science and engineering, with CEFR proficiency bands ranging from B1 to C1. The power specification targeted the condition-by-time interaction in the primary outcome. Under $\alpha = .05$, 80% power, a standardized interaction of approximately 0.30, a repeated-measures correlation of approximately .62, modest site and instructor variation, 12 delivery groups, and 10% attrition, 252 completers were required [39]. The recruited sample supported the primary effect but provided lower power for moderator interactions; therefore, moderation was treated as a secondary confirmatory analysis.

3. ETHICS APPROVAL, CONSENT, AND PRIVACY PROTECTION

The Research Ethics Committee of the coordinating university approved the study on 28 January 2021. Participation was voluntary, was unrelated to course grades, and could be discontinued without penalty. Before enrolment, students received an information sheet describing the study aims, data requirements, privacy risks, withdrawal procedures, and data-sharing limits. Separate written consent was obtained for digital-text submission, controlled writing tasks, questionnaire responses, interview recording, and the use of deidentified quotations.

Participants reviewed their message samples before submission and could remove any item without explanation. We removed names, usernames, telephone numbers, email addresses, dates, locations, institutional identifiers, hyperlinks, and third-party details before analysis. Only participant-authored turns were retained. We assigned each participant a study code and stored the identity key separately from the research files. Encrypted data were held on a restricted university server accessible only to the principal investigator and data manager. The team deleted audio files after transcript verification and retained deidentified analytical files for five years, in accordance with the approved data-management plan.

4. ALLOCATION, INTERVENTION CONTROL, AND FIDELITY

After baseline assessment, an independent statistician prepared a computerized 1:1 allocation sequence using variable block sizes of four and six, stratified by site, proficiency band, and study level. The sequence was stored in an encrypted file inaccessible to recruiters and was released only after baseline completion. Twelve delivery groups were formed, six per condition, with 20-27 students per group. Six instructors each taught one intervention and one comparison group at the same site, and group order was balanced to reduce instructor-by-condition confounding. The primary models included random intercepts for participant, instructional group, and instructor; site was treated as a fixed effect because only three purposively selected sites were included. Writing raters and the primary analyst received masked condition and wave labels.

The intervention comprised eight weekly 75-minute sessions. Instructor training lasted six hours and used standardized lesson plans, model texts, and fidelity checklists. The intervention and comparison groups received the same number and duration of sessions, and the comparison condition completed parallel academic-writing activities without explicit digital-to-formal contrastive instruction. Mean attendance was 6.9 sessions in the intervention arm and 6.8 in the comparison arm; 82.5% attended at least six sessions. Independent fidelity review indicated that 89.2% of intervention components and 91.0% of comparison-condition components were delivered as scheduled. We recorded two brief cross-condition discussions and three timetable-related deviations and examined them in sensitivity analyses.

Table 3. Intervention sequence and fidelity evidence.

Week	Focus	Core activity	Fidelity evidence
1	Register and audience	Compare digital, academic, and professional samples	Context-identification record; lesson checklist
2	Spelling and capitalization	Expand textisms and revise capitalization	Edited contrastive text; completion log
3	Punctuation and sentence closure	Diagnose fragments, run-ons, and expressive punctuation	Boundary-revision exercise; instructor log
4	Grammar and contractions	Compare conversational compression with formal clauses	Formality justification; learner explanation
5	Lexical choice and tone	Revise vague or overly informal vocabulary	Audience-specific vocabulary map
6	Cohesion and coherence	Expand shared context into explicit links	Reconstructed paragraph; cohesion task
7	Proofreading routines	Apply focused editing passes	Revision log; attendance record
8	Integrated register shifting	Transform a message thread into a formal email	Portfolio task; fidelity review

5. DIGITAL-MESSAGE SAMPLING AND WRITING TASKS

Participants submitted 500-800 English words written during the preceding 30 days and distributed across at least three days, two recipient categories, and two communicative purposes. Sampling instructions requested ordinary messages rather than unusually polished or unusually informal examples. We recorded platform, date band, recipient category, purpose, autocorrection, predictive-text use, and exclusions. The median submitted sample contained 668 participant-authored words.

At each wave, every participant completed a 600-800-word argumentative essay in 70 minutes and a 250-350-word formal email in 30 minutes. Essay and email outcomes were treated as separate genres. Parallel prompts were piloted with 24 students who did not enter the main study. Dictionaries, spell-checkers, grammar checkers, predictive text, and external composition software were unavailable during the controlled tasks except where an approved accommodation applied. Administration time, interruptions, accommodations, and deviations were logged.

6. INSTRUMENTS, CODING, AND RELIABILITY

Digital features were coded per 1,000 words or per eligible opportunity, with document prevalence reported for sparse forms. The feature dictionary distinguished textisms and abbreviations from ordinary lexical uses, and it did not classify expressive digital punctuation as an error unless it violated the conventions of the formal task. We rated formal texts on six-point analytic scales for grammar, spelling and

mechanics, lexical resources, syntactic control, cohesion and coherence, register appropriateness, development, and task response.

English proficiency was measured with the Oxford Online Placement Test (OOPT), and mapped to B1, B2, or C1 using the publisher's CEFR interpretation. Register awareness was assessed through an 18-item detection, rewriting, and explanation task. We measured proofreading behavior with a six-item scale and a post-task revision log. The digital-writing questionnaire contained 22 items addressing frequency, platform, recipient, purpose, language background, autocorrection, predictive text, and revision habits. Internal consistency was $\omega = .87$ for register awareness and $\alpha = .84$ for proofreading.

Three qualified writing specialists completed a two-hour training and calibration session using texts excluded from the analysis. Two raters independently rated every formal text, and 25% received a third rating to monitor drift. Raters were blinded to condition, site, and wave. ICC(2,k) for 1,068 double-rated texts was .95, 95% CI [.94, .96], and quadratic weighted κ was .76, 95% CI [.73, .79]. Dictionary-based feature coding was validated against 600 manually annotated texts; feature-level F1 values ranged from .71 to .93. Because punctuation and spelling precision were .60 and .62, respectively, the focal models were repeated with human-verified cases, high-confidence thresholds, and regression calibration for misclassification. The focal coefficients changed by less than 0.04 standardized units. Disagreements exceeding one scale point or one coding category were adjudicated by the senior rater.

7. QUALITATIVE SAMPLING AND ANALYSIS

After follow-up, 36 participants were selected through maximum variation across condition, site, proficiency, digital-feature density, register-awareness score, and outcome change. Interviews were conducted between 5 and 16 July 2021, lasted 29-52 minutes, and combined semi-structured questions with stimulated recall based on participants' own writing. We audio-recorded interviews, transcribed them verbatim, checked them against the recordings, and assigned transcript identifiers.

Framework analysis followed the procedure in [29] and proceeded through familiarization, initial coding, matrix development, indexing, charting, and cross-case interpretation. The coding framework covered audience recognition, automaticity, digital identity, autocorrection, time pressure, revision, and response to instruction. A second analyst independently reviewed nine transcripts and the full joint-display matrix; the team resolved disagreements through discussion and documented them in the audit log. We retained negative and divergent cases to challenge overly simple interpretations. Quotations were reported only when the participant had separately consented to quotation use. Reporting followed established journal standards for qualitative and mixed-methods research [40].

IX. STATISTICAL ANALYSIS

We summarized feature distributions with means, medians, standard deviations, interquartile ranges, zero proportions, document prevalence, and 95% confidence intervals. Paired register contrasts used paired t tests or Wilcoxon signed-rank tests, depending on distributional properties, with standardized paired effects, bootstrap confidence intervals, and false-discovery-rate control within prespecified feature families [41].

RQ2 and RQ3 used baseline predictors and baseline formal outcomes. We analyzed comparable-form counts with negative-binomial mixed models, using eligible formal words as an offset. Overall quality and other approximately continuous outcomes were analysed with linear mixed models. Fixed effects included prompt and site, while random intercepts accounted for participant, instructional group, and instructor. We standardized continuous predictors within the full baseline sample, and we standardized outcomes using the pooled baseline standard deviation. We conducted ordinal mixed-model sensitivity analyses for the six-point outcomes. We reported model-specific N, unstandardized or standardized estimates, marginal and conditional R^2, random-effect estimates, dispersion, residual diagnostics, multicollinearity, and influential-case checks.

Confirmatory moderation was restricted to digital density × register awareness and digital density × proofreading. We reported confidence intervals, multiplicity-adjusted inference, and simple slopes at low,

mean, and high moderator values. We treated age, first language, discipline, platform, reading frequency, and writing-support-tool use as exploratory variables.

The primary intervention analysis followed the intention-to-treat principle and included all 280 randomized participants. Missingness ranged from 0.0% for allocation to 9.3% for follow-up outcomes and 3.2% for covariates. Multiple imputation by chained equations used 30 datasets and 20 iterations, with pooled estimates, trace plots, and distributional convergence checks [42]. Per-protocol analysis required attendance at six or more sessions and no major protocol violation. We conducted complete-case, ordinal mixed-model, human-verified coding, high-confidence coding, regression-calibrated misclassification, instructor/group-adjusted, contamination-excluded, and age-adjusted analyses as sensitivity checks.

X. RESULTS

1. PARTICIPANT FLOW AND BASELINE CHARACTERISTICS

The recruitment process yielded 360 invitations, 322 screening records, 294 eligible students, and 280 randomized participants. Each condition included 140 participants, and 127 participants per condition completed follow-up. The primary intention-to-treat models included all 280 participants, while complete-case follow-up descriptions used 254 participants. Baseline balance was acceptable for most reported characteristics, but age showed an absolute standardized mean difference of 0.25 and therefore exceeded the prespecified descriptive threshold of 0.15. All other calculable standardized differences were below 0.15. Age was retained as an exploratory covariate, and age-adjusted sensitivity analysis did not change the direction or practical interpretation of the focal effects.

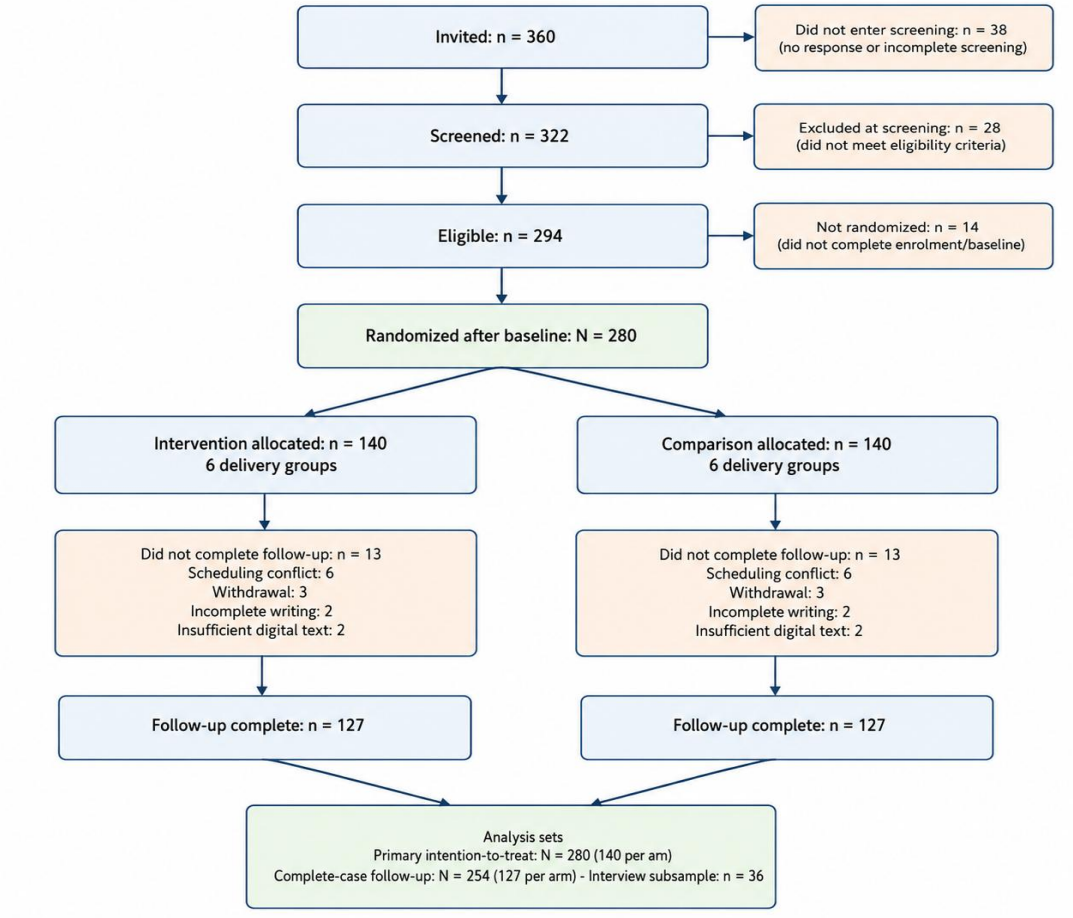


FIGURE 2. CONSORT-style participant flow, exclusions, arm-specific attrition reasons, and analysis sets.

Table 4. Participant characteristics by condition.

Characteristic	Intervention (n=140)	Comparison (n=140)	Total (N=280)	Absolute SMD
Age, mean (SD)	22.2 (3.0)	23.0 (3.3)	22.6 (3.2)	0.25
Women, n (%)	77 (55.0)	78 (55.7)	155 (55.4)	0.01
Undergraduate, n (%)	102 (72.9)	103 (73.6)	205 (73.2)	0.02
B1 proficiency, n (%)	43 (30.7)	43 (30.7)	86 (30.7)	0.00
B2 proficiency, n (%)	66 (47.1)	67 (47.9)	133 (47.5)	0.01
C1 proficiency, n (%)	31 (22.1)	30 (21.4)	61 (21.8)	0.02
Digital messages/day, median [IQR]	48 [38-66]	50 [37-68]	50 [37-67]	—
Baseline overall quality, mean (SD)	3.54 (0.66)	3.60 (0.75)	3.57 (0.71)	0.09
Register awareness, mean (SD)	17.68 (3.36)	17.20 (3.56)	17.44 (3.46)	0.14
Proofreading, mean (SD)	3.50 (0.64)	3.47 (0.70)	3.48 (0.67)	0.04

2. MEASUREMENT QUALITY

Average analytic scores showed excellent inter-rater reliability, ICC(2,k) = .95, 95% CI [.94, .96], and rounded ordinal categories showed substantial agreement, quadratic weighted κ = .76, 95% CI [.73, .79]. Register-awareness and proofreading scales showed acceptable internal consistency. Dictionary-based feature classification was strongest for textisms and weakest for punctuation and spelling; we addressed the lower precision of these two categories with human-verified, high-confidence, misclassification-corrected sensitivity analyses.

Table 5. Reliability and feature-validation results.

Measure	Validation sample	Statistic	Estimate	Interpretation
Formal-writing analytic scores	1,068 double-rated texts	ICC(2,k)	0.95 [0.94, 0.96]	Excellent agreement
Rounded rating categories	1,068 double-rated texts	Quadratic weighted κ	0.76 [0.73, 0.79]	Substantial agreement
Register-awareness task	N = 280	McDonald's ω	0.87	Good internal consistency
Proofreading scale	N = 280	Cronbach's α	0.84	Good internal consistency
Capitalization classification	600 manually annotated texts	Precision / recall / F1	0.76 / 0.96 / 0.85	Acceptable-to-strong validity
Grammar classification	600 manually annotated texts	Precision / recall / F1	0.71 / 0.85 / 0.77	Acceptable validity

Measure	Validation sample	Statistic	Estimate	Interpretation
Punctuation classification	600 manually annotated texts	Precision / recall / F1	0.60 / 0.86 / 0.71	Limited precision; sensitivity checked
Spelling classification	600 manually annotated texts	Precision / recall / F1	0.62 / 0.84 / 0.71	Limited precision; sensitivity checked
Textism classification	600 manually annotated texts	Precision / recall / F1	0.96 / 0.89 / 0.93	Strong validity

3. CROSS-REGISTER DIFFERENCES

Digital texts contained substantially more textisms, contractions, repeated punctuation, lower-case pronoun i, and omitted final punctuation than essays or formal emails. Essays showed the highest lexical-sophistication index. The magnitude of the paired effects indicated strong register differentiation within participants rather than a uniform displacement of formal conventions.

Table 6. Baseline cross-register descriptive statistics.

Feature	Digital M (SD)	Essay M (SD)	Formal email M (SD)	Median digital	Paired effect d_z
Textisms / 1,000 words	45.51 (17.21)	2.34 (0.92)	3.36 (1.21)	45.64	2.58
Contractions / 1,000 words	65.40 (15.59)	14.45 (4.80)	30.20 (7.58)	65.91	3.18
Repeated punctuation / 1,000 characters	13.43 (4.55)	1.11 (0.38)	1.76 (0.55)	13.60	2.70
Lowercase i / 1,000 words	5.52 (2.29)	0.55 (0.28)	0.94 (0.38)	5.53	2.16
Omitted final punctuation, %	31.73 (10.28)	3.08 (1.56)	5.53 (1.95)	32.24	2.72
Lexical sophistication	0.17 (0.03)	0.26 (0.04)	0.23 (0.03)	0.17	-1.67

4. PARTICIPANT-LEVEL ASSOCIATIONS AND MODERATION

In the baseline negative-binomial model (N = 280; marginal R^2 = .29; conditional R^2 = .41; dispersion ratio = 1.08), a one-standard-deviation increase in digital-feature density was associated with a 27.7% higher expected rate of comparable formal forms (RR = 1.28, 95% CI [1.11, 1.46], $p < .001$). In the baseline quality model (N = 280; marginal R^2 = .52; conditional R^2 = .58), digital density was not independently associated with overall writing quality ($\beta = 0.05$, 95% CI [-0.03, 0.12], $p = .203$). Register awareness, proofreading, and proficiency were associated with lower comparable-form rates and with higher overall quality. Random-intercept standard deviations in the count model were 0.18 for instructional group and 0.07 for instructor.

The digital density × register awareness interaction remained significant after false-discovery-rate adjustment (RR = 0.85, 95% CI [0.75, 0.97], adjusted p = .035). Marginal rate ratios at low, mean, and high awareness were 1.49, 1.28, and 1.09, respectively. The digital density × proofreading interaction was directionally protective but did not reach the adjusted threshold (RR = 0.91, 95% CI [0.81, 1.03], adjusted p =

.125); corresponding marginal rate ratios were 1.38, 1.28, and 1.18. The interaction patterns were associational and did not establish a causal regulation mechanism.

Table 7. Association models.

Outcome	Predictor	Effect estimate	95% CI	p	N	Marginal / conditional R^2
Comparable-form count	Digital-feature density	RR 1.28	[1.11, 1.46]	< .001	280	.29 / .41
Comparable-form count	Register awareness	RR 0.61	[0.54, 0.69]	< .001	280	.29 / .41
Comparable-form count	Proofreading	RR 0.81	[0.72, 0.91]	< .001	280	.29 / .41
Comparable-form count	Proficiency	RR 0.80	[0.71, 0.90]	< .001	280	.29 / .41
Overall quality	Digital-feature density	β = 0.05	[−0.03, 0.12]	.203	280	.52 / .58
Overall quality	Register awareness	β = 0.43	[0.36, 0.50]	< .001	280	.52 / .58
Overall quality	Proofreading	β = 0.24	[0.17, 0.31]	< .001	280	.52 / .58
Overall quality	Proficiency	β = 0.65	[0.58, 0.72]	< .001	280	.52 / .58

Model diagnostics. The count-model dispersion ratio was 1.08; a zero-inflated alternative did not materially improve fit. The maximum variance-inflation factor was 2.1 and the maximum Cook's distance was .04. Residual, leverage, influence, convergence, and singularity checks did not identify a result-changing case. Group and instructor random effects were retained. Human-verified, high-confidence, and regression-calibrated sensitivity models changed the focal coefficients by less than 0.04 standardized units.

Table 8. Confirmatory moderation tests.

Interaction	Rate ratio	95% CI	Raw p	FDR-adjusted p	Predicted RR: low / mean / high	Conclusion
Digital density × register awareness	0.85	[0.75, 0.97]	.018	.035	1.49 / 1.28 / 1.09	Supported
Digital density × proofreading	0.91	[0.81, 1.03]	.125	.125	1.38 / 1.28 / 1.18	Not supported

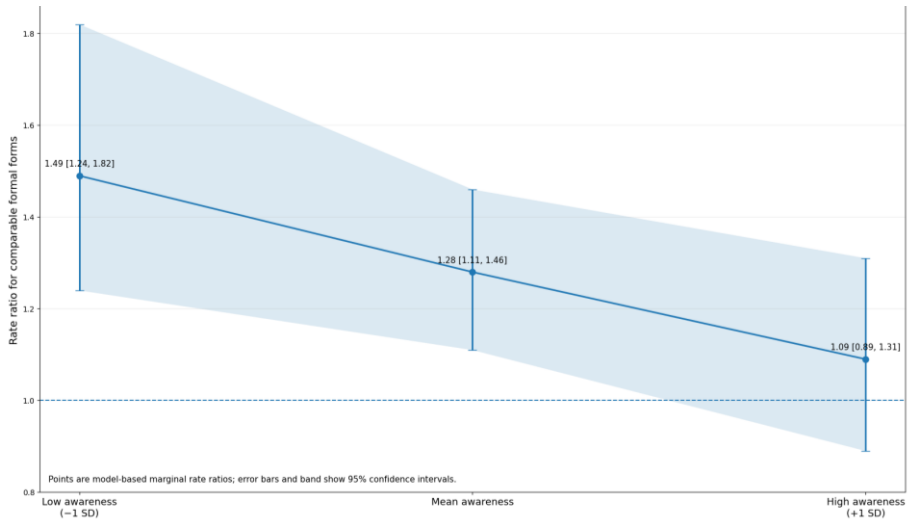


FIGURE 3. Moderating role of register awareness in the digital-density association, with model-based 95% confidence intervals.

5. INTERVENTION OUTCOMES

The pooled intention-to-treat models (N = 280) showed larger baseline-to-follow-up gains in the intervention condition for grammar, spelling and mechanics, register appropriateness, and overall quality. The largest standardized condition-by-time effect concerned register appropriateness, while the overall-quality effect was small. Comparable digital-associated forms also declined more strongly in the intervention condition. Per-protocol, complete-case, ordinal mixed-model, instructor/group-adjusted, contamination-excluded, human-verified coding, high-confidence coding, and misclassification-corrected estimates were similar in direction and magnitude; the register-appropriateness effect ranged from 0.45 to 0.51 across sensitivity analyses.

Table 9. Baseline and follow-up outcomes with adjusted intention-to-treat effects (N = 280).

Outcome	Intervention baseline	Intervention follow-up	Comparison baseline	Comparison follow-up	Effect [95% CI]	Magnitude
Grammar	3.54 (0.59)	3.76 (0.67)	3.50 (0.67)	3.60 (0.73)	0.18 [0.07, 0.29], p = .001	Small
Spelling/mechanics	3.50 (0.61)	3.80 (0.74)	3.55 (0.70)	3.61 (0.71)	0.33 [0.23, 0.44], p < .001	Small-to-moderate
Register appropriateness	3.48 (0.62)	3.89 (0.69)	3.44 (0.67)	3.48 (0.73)	0.48 [0.39, 0.58], p < .001	Moderate
Overall quality	3.54 (0.66)	3.75 (0.75)	3.60 (0.75)	3.67 (0.77)	0.19 [0.10, 0.28], p < .001	Small
Comparable forms / 1,000 words	2.01	0.81	1.73	1.71	RR 0.41 [0.27, 0.62], p < .001	Small-to-moderate decline

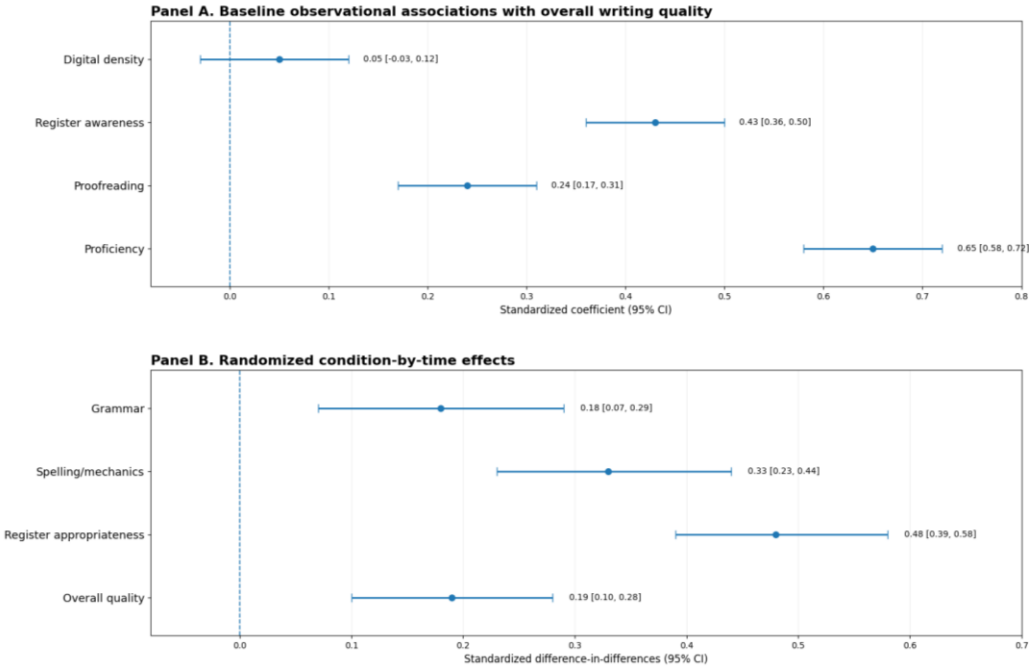


FIGURE 4. Panel A: baseline observational associations with overall writing quality. Panel B: randomized instructional condition-by-time effects. Error bars show 95% confidence intervals.

6. QUALITATIVE FINDINGS AND MIXED-METHODS INTEGRATION

Framework analysis identified four principal patterns. First, participants with high register-awareness scores and strong formal performance often shifted their writing to suit the audience. Second, automaticity under time pressure accounted for some similar formal occurrences among high-density digital writers. Third, intervention responders described proofreading as audience redesign rather than simple error correction. Fourth, negative cases showed that weak development and cohesion could occur even when digital-associated forms were rare.

Audience-sensitive shifting was represented in 31 interviews, proofreading as audience redesign in 27, automaticity under time pressure in 22, and digital identity with selective formality in 19. Seven participants provided negative or divergent cases that challenged a simple deficit explanation. The joint display showed convergence between high awareness and successful register shifting, complementarity between time-pressure accounts and count-model results, and divergence when low-quality writing occurred without textisms.

Table 10. Mixed-methods joint display.

Case profile	Quantitative evidence	Theme	Extract	Evidence ID	Integration
Low digital density / weak quality	Density 29.5; baseline quality 2.64; awareness 19.6	Negative case: poor quality without textisms	My problem was not texting language. I struggled to organize the argument and connect the paragraphs even	TR01, lines 44-48	Divergence

Case profile	Quantitative evidence	Theme	Extract	Evidence ID	Integration
			when I used standard spelling.		
High awareness / strong formality	Density 25.0; baseline quality 3.60; awareness 18.9	Digital identity and selective formality	Informal spelling feels natural with close friends, but I do not see it as the same kind of writing as a report.	TR02, lines 61-64	Convergence
Intervention responder	Density 36.7; quality change +0.74; awareness 19.6	Proofreading as audience redesign	The checklist changed my revision. I stopped reading only for ideas and started reading again for audience, tone, and sentence endings.	TR05, lines 72-77	Convergence
High density / lower awareness	Density 44.9; baseline quality 3.43; awareness 13.1	Automaticity under time pressure	When I am rushing, the short form appears before I notice it. I usually catch it only if I make a separate editing pass.	TR06, lines 36-40	Complementarity
High density / successful shifting	Density 59.0; baseline quality 3.05; awareness 18.0	Audience-sensitive shifting	I write quickly with friends, but when the reader is a lecturer, I expand the words and check how the message sounds.	TR14, lines 51-55	Convergence

XI. DISCUSSION`

The results supported strong register differentiation rather than a generalized linguistic-decline account. Digital messages contained markedly more abbreviations, contractions, expressive punctuation, lower-case forms, and omitted sentence closure than academic and professional texts. The magnitude of the paired contrasts showed that the same writers maintained substantially different linguistic profiles across contexts, consistent with research on texting, school writing, and spelling [3]–[5]. Related work on Spanish textisms and perceptions of textese provides additional support [6]–[8].

The participant-level count model identified a modest positive association between digital-feature density and comparable forms in formal writing, but digital density did not explain overall quality once

proficiency, register awareness, and proofreading were considered. This distinction is theoretically important. Surface-form accessibility can coexist with strong formal writing, whereas broader quality depends on sentence control, lexical resources, and cohesion [17]–[19]. Accuracy, discourse development, and syntactic complexity further contribute to writing quality [20]–[22]. Corpus-derived syntactic measures provide complementary evidence [23].

Register awareness significantly attenuated the digital-density association. At low awareness, digital density was more strongly related to comparable formal forms; at high awareness, the association was substantially weaker. Proofreading was independently protective, although its interaction with digital density was not significant after adjustment. The findings therefore support a register-mediated account in which contextual recognition and revision regulate the formal expression of frequently available forms, but they do not establish a complete causal mechanism.

The intervention produced the largest improvement in register appropriateness, followed by spelling, mechanics, and grammar. The global-quality effect was small. This hierarchy matched the instructional focus: the program directly targeted audience, tone, sentence closure, capitalization, contraction choice, and focused proofreading, whereas broad rhetorical development normally depends on longer-term growth in reading, disciplinary knowledge, vocabulary, argumentation, and genre experience.

The qualitative evidence clarified the quantitative pattern. Participants distinguished readers and purposes, described automatic short forms under time pressure, and reported using proofreading to redesign messages for formal audiences. Negative cases were especially important because they showed that the absence of textisms did not guarantee strong writing. Several participants with low digital-feature density still produced weak development, cohesion, or task response, reinforcing the distinction between surface transfer and multidimensional quality.

Alternative explanations remained relevant. Voluntary message selection may have privileged safer or unusually careful examples, and autocorrection may have both prevented and introduced errors. Participants and instructors could not be blinded, creating possible expectancy and instructor effects. Equal contact time, independent fidelity review, masked rating, and site-adjusted analysis reduced these threats but did not eliminate them. The similarity of intention-to-treat and sensitivity estimates strengthened confidence in the main intervention pattern, but the short follow-up prevented conclusions about long-term retention.

The study's international and multilingual relevance lies in treating digital variation as a contextual resource rather than a unitary deficit. Many university writers use English alongside other languages and shift across interpersonal, academic, and professional settings. Instruction that makes audience, genre, and revision decisions explicit may therefore be more productive than blanket prohibition of digital forms. At the same time, researchers should not generalize the findings beyond comparable university contexts without further cross-site and cross-linguistic replication.

1. THEORETICAL AND METHODOLOGICAL CONTRIBUTIONS

The findings supported a register-mediated account that distinguished digital-form availability, contextual selection, multidimensional writing quality, and instructional change. Register theory explained the strong within-participant differentiation across digital, academic, and professional tasks, while writing-process theory accounted for the role of deliberate review and revision.

Methodologically, the study paired registers within writers, limited confirmatory moderators, used count-appropriate and mixed-effects models, separated observational associations from instructional effects, and linked qualitative claims to transcript identifiers. These features strengthened construct validity and reduced the risk of interpreting a small number of visible textisms as evidence of generalized writing decline.

XII. PRACTICAL AND PEDAGOGICAL IMPLICATIONS

The findings suggest potential value in contrastive register instruction, digital-to-formal rewriting, audience-based tone analysis, and focused proofreading passes. These practices may support contextual control by helping students identify which features are effective in informal interaction and which require expansion or revision in academic and professional genres.

The moderate effect on register appropriateness and smaller effect on overall quality indicate that a short intervention should not be presented as a complete solution for academic writing development. Broader gains require sustained work on reading, argumentation, disciplinary knowledge, lexical resources, cohesion, and genre-specific rhetorical practice. Register-focused activities are therefore best integrated into a wider writing curriculum rather than offered as an isolated remedial program.

XIII. LIMITATIONS

The study involved three public universities within one geographic region, which limits generalization to institutions with different linguistic, curricular, and technological environments. Although the sample included several disciplines and CEFR bands, it did not represent all university populations or age groups.

Digital-message submission was voluntary. Participants could remove sensitive exchanges, and privacy safeguards required deleting third-party and contextual information. The resulting samples may therefore have underrepresented highly personal, emotionally charged, or unusually informal communication. Coverage indicators reduced this bias but could not eliminate it.

Several measures relied partly on self-report, including message frequency, proofreading habits, autocorrection, and predictive-text use. These responses may have been affected by recall error or social desirability. The register-awareness task and proofreading scale showed acceptable reliability, but they captured selected aspects of contextual control rather than the full complexity of real-time writing decisions.

Participants and instructors could not be blinded to the instructional condition. Expectancy effects, instructor enthusiasm, and limited contamination between groups may have contributed to change. Equal contact time, fidelity monitoring, masked rating, and sensitivity analyses reduced these risks but did not eliminate them.

The eight-week intervention and immediate follow-up were sufficient to detect short-term change but not durable learning. A delayed retention assessment would be necessary to determine whether improvements in register appropriateness and mechanics persisted. The sample was also powered for the primary condition-by-time effect rather than numerous subgroup interactions, so interpret exploratory differences by first language, discipline, platform, or age cautiously.

Dictionary-based feature extraction performed adequately overall, but punctuation and spelling classification had limited precision (.60 and .62). Human verification, high-confidence thresholds, adjudication, and regression-calibrated sensitivity analyses reduced the risk of measurement-driven conclusions; the focal coefficients changed by less than 0.04 standardized units. Some context-sensitive cases nevertheless remained difficult to classify, and future studies should extend the feature dictionary and examine additional revision-process evidence, including keystroke logging.

XIV. CONCLUSION

This paired longitudinal mixed-methods study examined how university students' digital-writing practices related to grammar, spelling, register appropriateness, and formal-writing quality and evaluated an eight-week register-shifting and proofreading intervention. Digital messages and formal texts showed strong within-writer differentiation. Higher digital-feature density was associated with a modest increase in comparable formal forms, but it was not independently related to overall writing quality after proficiency, register awareness, and proofreading were considered. Register awareness reduced the strength of the digital-density association, supporting a conditional rather than deterministic account of influence. The intervention produced its largest gain in register appropriateness, with smaller improvements in mechanics, grammar, and overall quality. Interview evidence further showed that audience recognition, time pressure, identity, and deliberate revision shaped movement across registers. These findings do not support a generalized decline account. They instead indicate that digital and formal writing are related through selective surface-form accessibility and contextual regulation. Register-focused instruction may strengthen formal control, but broader writing development requires sustained attention to argumentation, disciplinary knowledge, lexical development, cohesion, and genre.

Author Contributions

The author conducted the conceptualization, methodology, data analysis, investigation, writing, review, editing, and final approval of the manuscript.

Funding

This research received no external funding.

Data Availability

The dataset will be available from the author upon reasonable request.

Conflicts of Interest

The author declares no conflict of interest.

REFERENCES

[1]. T. Chen and M.-Y. Kan, "Creating a live, public short message service corpus: The NUS SMS Corpus," Language Resources and Evaluation, vol. 47, no. 2, pp. 299–335, 2013, doi: 10.1007/s10579-012-9197-9.

[2]. F. Lyddy, F. Farina, J. Hanney, L. Farrell, and N. Kelly O'Neill, "An analysis of language in university students' text messages," Journal of Computer-Mediated Communication, vol. 19, no. 3, pp. 546–561, 2014, doi: 10.1111/jcc4.12045.

[3]. C. N. van Dijk, M. van Witteloostuijn, N. Vasić, S. Avrutin, and E. Blom, "The influence of texting language on grammar and executive functions in primary school children," PLOS ONE, vol. 11, no. 3, e0152409, 2016, doi: 10.1371/journal.pone.0152409.

[4]. L. Verheijen, W. Spooren, and A. van Kemenade, "Relationships between Dutch youths' social media use and school writing," Computers and Composition, vol. 56, art. 102574, 2020, doi: 10.1016/j.compcom.2020.102574.

[5]. L. Verheijen and W. Spooren, "The impact of WhatsApp on Dutch youths' school writing and spelling," Journal of Writing Research, vol. 13, no. 1, pp. 155–191, 2021, doi: 10.17239/jowr-2021.13.01.05.

[6]. A. Gómez-Camacho, C. I. Hunt-Gómez, and A. Valverde-Macías, "Textisms, texting, and spelling in Spanish," Lingua, vol. 201, pp. 92–101, 2018, doi: 10.1016/j.lingua.2017.09.004.

[7]. N. Kemp and A. Grace, "Txting across time: Undergraduates' use of textese in seven consecutive first-year psychology cohorts," Writing Systems Research, vol. 9, no. 1, pp. 82–98, 2017, doi: 10.1080/17586801.2017.1285220.

[8]. N. Kemp, J. Graham, R. Grieve, and E. Beyersmann, "The influence of textese on adolescents' perceptions of text message writers," Telematics and Informatics, vol. 65, 101720, 2021, doi: 10.1016/j.tele.2021.101720.

[9]. J. Parrella, H. R. Leggette, and T. Redwine, "Measuring the correlation between digital media usage and students' perceived writing ability: Are they related?" Research in Learning Technology, vol. 29, 2506, 2021, doi: 10.25304/rlt.v29.2506.

[10]. I. Elola and A. Oskoz, "Writing with 21st-century social tools in the L2 classroom: New literacies, genres, and writing practices," Journal of Second Language Writing, vol. 36, pp. 52–60, 2017, doi: 10.1016/j.jslw.2017.04.002.

[11]. M. Li, "Computer-mediated collaborative writing in L2 contexts: An analysis of empirical research," Computer Assisted Language Learning, vol. 31, no. 8, pp. 882–904, 2018, doi: 10.1080/09588221.2018.1465981.

[12]. C. Strobl, E. Ailhaud, K. Benetos, A. Devitt, O. Kruse, A. Proske, and C. Rapp, "Digital support for academic writing: A review of technologies and pedagogies," Computers & Education, vol. 131, pp. 33–48, 2019, doi: 10.1016/j.compedu.2018.12.005.

[13]. C. A. Hafner and W. Y. J. Ho, "Assessing digital multimodal composing in second language writing: Towards a process-based model," Journal of Second Language Writing, vol. 47, 100710, 2020, doi: 10.1016/j.jslw.2020.100710.

[14]. L. Jiang, "The affordances of digital multimodal composing for EFL learning," ELT Journal, vol. 71, no. 4, pp. 413–422, 2017, doi: 10.1093/elt/ccw098.

[15]. L. Jiang, M. Yang, and S. Yu, "Chinese ethnic minority students' investment in English learning empowered by digital multimodal composing," TESOL Quarterly, vol. 54, no. 4, pp. 954–979, 2020, doi: 10.1002/tesq.566.

[16]. D. Biber and S. Conrad, Register, Genre, and Style, 2nd ed. Cambridge, UK: Cambridge University Press, 2019, doi: 10.1017/9781108686136.

[17]. S. A. Crossley, "Linguistic features in writing quality and development: An overview," Journal of Writing Research, vol. 11, no. 3, pp. 415–443, 2020, doi: 10.17239/jowr-2020.11.03.01.

[18]. K. Kyle and S. A. Crossley, "The relationship between lexical sophistication and independent and source-based writing," Journal of Second Language Writing, vol. 34, pp. 12–24, 2016, doi: 10.1016/j.jslw.2016.10.003.

[19]. M. Kim, S. A. Crossley, and K. Kyle, "Lexical sophistication as a multidimensional phenomenon: Relations to second language lexical proficiency, development, and writing quality," The Modern Language Journal, vol. 102, suppl. 1, pp. 120–141, 2018, doi: 10.1111/modl.12447.

[20]. C. Polio and M. C. Shea, "An investigation into current measures of linguistic accuracy in second language writing research," Journal of Second Language Writing, vol. 26, pp. 10–27, 2014, doi: 10.1016/j.jslw.2014.09.003.
[21]. A. C. Graesser, D. S. McNamara, M. M. Louwerse, and Z. Cai, "Coh-Metrix: Analysis of text on cohesion and language," *Behavior Research Methods, Instruments, & Computers*, vol. 36, no. 2, pp. 193–202, 2004, doi: 10.3758/BF03195564.
[22]. S. A. Crossley and D. S. McNamara, "Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners," Journal of Second Language Writing, vol. 26, pp. 66–79, 2014, doi: 10.1016/j.jslw.2014.09.006.
[23]. X. Lu, "A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development," *TESOL Quarterly*, vol. 45, no. 1, pp. 36–62, 2011, doi: 10.5054/tq.2011.240859.
[24]. D. Bikowski and R. Vithanage, "Effects of web-based collaborative writing on individual L2 writing development," Language Learning & Technology, vol. 20, no. 1, pp. 79–99, 2016, doi: 10.64152/10125/44447.
[25]. [25] R. Negretti and L. McGrath, "Scaffolding genre knowledge and metacognition: Insights from an L2 doctoral research writing course," Journal of Second Language Writing, vol. 40, pp. 12–31, 2018, doi: 10.1016/j.jslw.2017.12.002.
[26]. J. Bitchener and U. Knoch, "The contribution of written corrective feedback to language development: A ten-month investigation," *Applied Linguistics*, vol. 31, no. 2, pp. 193–214, 2010, doi: 10.1093/applin/amp016.
[27]. N. Shintani, R. Ellis, and W. Suzuki, "Effects of written feedback and revision on learners' accuracy in using two English grammatical structures," *Language Learning*, vol. 64, no. 1, pp. 103–131, 2014, doi: 10.1111/lang.12029.
[28]. J. R. Hayes, "Modeling and remodeling writing," Written Communication, vol. 29, no. 3, pp. 369–388, 2012, doi: 10.1177/0741088312451260.
[29]. N. K. Gale, G. Heath, E. Cameron, S. Rashid, and S. Redwood, "Using the framework method for the analysis of qualitative data in multi-disciplinary health research," *BMC Medical Research Methodology*, vol. 13, art. 117, 2013, doi: 10.1186/1471-2288-13-117.
[30]. L. S. Nowell, J. M. Norris, D. E. White, and N. J. Moules, "Thematic analysis: Striving to meet the trustworthiness criteria," *International Journal of Qualitative Methods*, vol. 16, no. 1, pp. 1–13, 2017, doi: 10.1177/1609406917733847.
[31]. D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," Journal of Statistical Software, vol. 67, no. 1, pp. 1–48, 2015, doi: 10.18637/jss.v067.i01.
[32]. A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest package: Tests in linear mixed effects models," Journal of Statistical Software, vol. 82, no. 13, pp. 1–26, 2017, doi: 10.18637/jss.v082.i13.
[33]. D. Lakens, "Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs," Frontiers in Psychology, vol. 4, art. 863, 2013, doi: 10.3389/fpsyg.2013.00863.
[34]. F. Faul, E. Erdfelder, A. Buchner, and A.-G. Lang, "Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses," Behavior Research Methods, vol. 41, no. 4, pp. 1149–1160, 2009, doi: 10.3758/BRM.41.4.1149.
[35]. M. L. McHugh, "Interrater reliability: The kappa statistic," Biochemia Medica, vol. 22, no. 3, pp. 276–282, 2012, doi: 10.11613/BM.2012.031.
[36]. M. D. Fetters, L. A. Curry, and J. W. Creswell, "Achieving integration in mixed methods designs—Principles and practices," Health Services Research, vol. 48, no. 6, pt. 2, pp. 2134–2156, 2013, doi: 10.1111/1475-6773.12117.
[37]. K. F. Schulz, D. G. Altman, and D. Moher, "CONSORT 2010 statement: Updated guidelines for reporting parallel group randomized trials," BMJ, vol. 340, c332, 2010, doi: 10.1136/bmj.c332.
[38]. M. E. Brooks et al., "glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling," The R Journal, vol. 9, no. 2, pp. 378–400, 2017, doi: 10.32614/RJ-2017-066.
[39]. M. Brysbaert and M. Stevens, "Power analysis and effect size in mixed effects models: A tutorial," Journal of Cognition, vol. 1, no. 1, art. 9, 2018, doi: 10.5334/joc.10.
[40]. H. M. Levitt et al., "Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The APA Publications and Communications Board task force report," American Psychologist, vol. 73, no. 1, pp. 26–46, 2018, doi: 10.1037/amp0000151.
[41]. Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," Journal of the Royal Statistical Society: Series B, vol. 57, no. 1, pp. 289–300, 1995, doi: 10.1111/j.2517-6161.1995.tb02031.x.
[42]. I. R. White, P. Royston, and A. M. Wood, "Multiple imputation using chained equations: Issues and guidance for practice," Statistics in Medicine, vol. 30, no. 4, pp. 377–399, 2011, doi: 10.1002/sim.4067.

https://doi.org/10.48161/ajst.v2n3a91

Appendix A. Essential Instruments, Scoring, and Reliability

This appendix summarizes the instruments required to interpret the reported constructs and outcomes. Full questionnaires, prompts, scoring keys, analytic rubrics, feature dictionaries, and administration instructions are provided in the Supplementary Materials.

Appendix Table A1. Essential instruments, administration, scoring, and reliability.

Instrument	Administration	Construct coverage	Scoring	Reliability / documentation
Digital-writing questionnaire	22 items; baseline and follow-up	Frequency, platform, recipient, purpose, language background, autocorrection, predictive text, and revision habits	Item-level variables and prespecified composites	Pilot review and item-to-construct mapping; full questionnaire in Supplementary Materials
Register-awareness task	18 detection, rewriting, and explanation items; administered before formal outcomes	Recognition and correction of register-incongruent language	0–24 points	McDonald's ω = .87; full items and scoring key in Supplementary Materials
Proofreading scale and revision log	Six-item scale plus post-task log at both waves	Focused rereading, mechanics, tone, sentence boundaries, formality, and final review	Mean score from 1 to 5; revision actions recorded separately	Cronbach's α = .84; full scale and log in Supplementary Materials
Academic essay	600–800 words; 70 minutes; parallel prompts at baseline and follow-up	Argumentation, grammar, mechanics, lexical resources, syntax, cohesion, development, and task response	Six-point analytic dimensions and composite overall-quality score	Double-rated; ICC(2,k) = .95; prompts and rubric in Supplementary Materials
Formal email	250-350 words; 30 minutes; parallel prompts at both waves	Audience, tone, register appropriateness, grammar, mechanics, cohesion, and task completion	Six-point analytic dimensions	Double-rated; ICC(2,k) = .95; prompts and rubric in Supplementary Materials
Oxford Online Placement Test (OOPT)	Supervised placement assessment administered at baseline	General English linguistic capacity	Publisher score mapped to B1, B2, or C1 CEFR band	Version, administration protocol, score report, and publisher CEFR interpretation retained with study records

Appendix B. Selected Qualitative Evidence and Negative Cases

The selected extracts below document thematic variation, intervention-related accounts, and negative cases. The Supplementary Materials provide the complete de-identified evidence matrix, coding framework, analyst memos, and transcript audit trail.

Appendix Table B1. Selected line-referenced qualitative evidence and analytical role.

Theme	Participant / transcript	Date	Lines	Extract	Analytical role
Audience-sensitive shifting	P214 / TR14	11 July 2021	51-55	I write quickly with friends, but when the reader is a lecturer, I expand the words and check how the message sounds.	Convergent evidence
Audience-sensitive shifting	P088 / TR09	9 July 2021	33-37	The greeting tells me what kind of language to use. With a lecturer, I slow down before I begin.	Convergent evidence
Automaticity under time pressure	P106 / TR06	7 July 2021	36-40	When I am rushing, the short form appears before I notice it. I usually catch it only if I make a separate editing pass.	Explanatory evidence
Automaticity under time pressure	P173 / TR22	13 July 2021	58-62	The mistake is more likely when I am trying to finish quickly, not because I do not know the correct form.	Explanatory evidence
Proofreading as audience redesign	P057 / TR05	7 July 2021	72-77	The checklist changed my revision. I reread for audience, tone, and sentence endings.	Intervention-linked evidence
Proofreading as audience redesign	P121 / TR18	12 July 2021	41-46	I used to correct only spelling. Now I check whether the sentence sounds suitable for the intended reader.	Intervention-linked evidence
Negative case	P148 / TR01	5 July 2021	44-48	My problem was not texting language. I struggled to organize the argument and	Negative or divergent evidence

Theme	Participant / transcript	Date	Lines	Extract	Analytical role
				connect the paragraphs even when I used standard spelling.	
Negative case	P021 / TR03	6 July 2021	39-43	I rarely use abbreviations, but my essay was still weak because I did not develop the examples.	Negative or divergent evidence

Note. Only selected, consented extracts are reproduced here; the complete de-identified evidence matrix and coding documentation are supplied as Supplementary Materials.