Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Cognitive Load, Writing Anxiety, and Syntactic Complexity in Teaching Turkish as a Foreign Language: An Experimental Intervention Study

**Kılıç Köçeri**

kilic.koceri@adalet.gov.tr

Ağrı İbrahim Çeçen University

**Ahmet Akçay**

Ağrı İbrahim Çeçen University

**Adnan Aydıntürk**

Ağrı İbrahim Çeçen University

**Akif Arslan**

Ağrı İbrahim Çeçen University

**Research Article**

**Keywords:** cognitive load theory, foreign language writing anxiety, syntactic complexity, teaching Turkish as a foreign language, intervention study

**Posted Date:** September 24th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11132963/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** No competing interests reported.

# Abstract

Writing in a foreign language ranks among the most cognitively demanding tasks confronting language learners, requiring the simultaneous management of linguistic form, affective states, and limited working memory capacity. Although growing evidence demonstrates that cognitive load (CL) and foreign language writing anxiety (FLWA) each independently suppress syntactic complexity (SC) in learner writing, no intervention study has yet systematically targeted both mechanisms within the Teaching Turkish as a Foreign Language (TaFL) context. This study employed a quasi-experimental pretest–posttest non-equivalent control group design with 300 international students across four universities in Türkiye. The 14-week Cognitive-Affective Writing Enhancement Programme (CAWEP) combined structured task design scaffolding, syntactic schema-building activities, and evidence-based anxiety regulation workshops. ANCOVA results revealed large, significant post-test advantages for the experimental group in extraneous CL ($\eta^2p = .27$), writing anxiety ($\eta^2p = .25$), and SC ($\eta^2p = .30$). Bootstrapped mediation analyses confirmed that FLWA partially mediates the extraneous CL–SC relationship, and concurrent think-aloud protocols provided convergent process-level evidence. Findings support integrating CL-informed task design and anxiety reduction into TaFL writing instruction.


# 1. Introduction

Writing in a second or foreign language (L2/FL) is widely acknowledged as the most cognitively demanding of the four language skills (Kormos, 2012; Porte & Richards, 2022). Unlike oral production, written composition requires the simultaneous coordination of orthographic, lexical, syntactic, discoursal, and pragmatic decisions under time constraints, all competing for limited working memory (WM) resources. This multi-layered burden is compounded in the Teaching Turkish as a Foreign Language (TaFL) context by Turkish agglutinative morphology, which places substantially higher processing demands on learners whose first-language typologies diverge from Turkish (Güneş, 2021; Yıldız & Yıldız, 2023). Despite these well-documented challenges, TaFL writing instruction remains an under-theorised domain, and pedagogically grounded intervention research is largely absent from the field.

Cognitive Load Theory (CLT; Sweller, 1988; Sweller et al., 2019) posits that human WM is limited in both capacity and duration, and that effective instructional design must manage three load components: intrinsic load (arising from task complexity), extraneous load (generated by suboptimal instructional design), and germane load (resources invested in schema formation). When extraneous load is elevated, cognitive resources are diverted from schema building, directly curtailing the depth of syntactic processing available to learners. A complementary body of scholarship demonstrates that FLWA (Cheng, 2004; Horwitz et al., 1986) exerts an independent negative effect on writing quality, particularly through syntactic avoidance: anxious writers restrict their use of complex structures to minimise the perceived risk of grammatical error (Woodrow, 2011; Zhan, 2024). Notwithstanding the theoretical coherence of these observations, the combined structural influence of CL and FLWA on SC, and the tractability of both through targeted intervention, remain empirically unexamined in TaFL contexts.

The present study addresses this lacuna through the design, implementation, and evaluation of the Cognitive-Affective Writing Enhancement Programme (CAWEP) a 14-week, theoretically grounded intervention targeting extraneous CL reduction, germane CL promotion, and FLWA regulation in TaFL writing classrooms. Employing a quasi-experimental pretest–posttest non-equivalent control group design with 300 participants, the study tests whether CAWEP produces measurable gains in SC relative to a business-as-usual control condition, while also examining the mediational mechanisms through which intervention effects operate. A concurrent think-aloud protocol strand with a 48-participant subsample provides process-level evidence illuminating the quantitative findings. The study contributes theoretically by integrating CLT and FLWA research traditions within an intervention framework, and practically by providing a replicable, evidence-based programme specification for TaFL writing instruction.

## 2. Literature Review

# 2.1. Cognitive Load Theory and Foreign Language Writing

Sweller's (1988) original formulation of CLT emerged from research on mathematical problem-solving and was subsequently extended to second language acquisition contexts (Leow, 2023; Sweller et al., 2019). The tripartite CL model distinguishes intrinsic, extraneous, and germane load based on their origins and functional relationships with learning outcomes. The element interactivity principle holds that intrinsic load derives not from the sheer number of task elements but from the degree to which they must be processed in an integrated manner a characterisation particularly apt for TaFL writing, where morphological, syntactic, and pragmatic decisions co-occur and interact (van Merriënboer & Sweller, 2022). Applied to FL writing, intrinsic load arises from the inherent complexity of the target language (especially for typologically distant L1–L2 pairs), extraneous load from poorly scaffolded tasks or ambiguous prompts, and germane load from the cognitive investment channelled toward constructing durable grammatical schemas. Révész et al. (2017) demonstrated that task complexity affects syntactic complexity and accuracy of L2 written output, providing indirect support for CLT predictions in writing contexts. More recently, Pham and Nguyen (2023) showed that structured pre-writing scaffolding reduces extraneous load and increases embedded clause use in EFL learner writing.

# 2.2. Foreign Language Writing Anxiety and Syntactic Complexity

FLWA is a trait-like affective construct that interacts with task features and instructional conditions to produce observable effects on writing quality (Cheng, 2004; Daly & Miller, 1975). Horwitz et al.'s (1986) FLCA model identifies communication apprehension, test anxiety, and fear of negative evaluation as central components; in FL writing, fear of negative evaluation is particularly salient because written products are more permanently evaluable than spoken output. Meta-analytic evidence consistently confirms the negative FLWA writing quality relationship. Huang (2018) reported a weighted mean correlation of r = − .38 across 34 studies; Zhan (2024) updated this estimate to r = − .41 (95% CI [− .48,

– .33]), with proficiency and task type as significant moderators. Most recently, Cheng et al. (2026) extended this line of inquiry with a longitudinal structural model showing that writing anxiety continues to exert a robust negative effect on L2 writing performance even after self-efficacy and self-regulated strategy use are statistically accounted for, and Lin (2026) reported that visually scaffolded pedagogical interventions can measurably reduce EFL writing anxiety alongside gains in learners' creative and syntactic elaboration, confirming that anxiety-focused intervention research in FL writing remains an active and rapidly developing area. In EFL contexts, FLWA has been specifically linked to reduced syntactic complexity (Woodrow, 2011) and shorter T-units (Bensoussan & Kreindler, 1990). In the TaFL domain, Ateş and Bahşi (2022) and Batmaz and Büyüikikiz (2025) documented significant negative associations between affective variables and writing performance.

## 2.3. Intervention Approaches in CL and Writing Anxiety Research

CLT-based instructional interventions consistently demonstrate that reducing extraneous load through improved task design enhances learning outcomes (van Merriënboer & Sweller, 2022). Key principles include worked examples, goal-free tasks, split-attention avoidance, and the completion effect (Sweller et al., 2019). Applied to L2 writing, these translate into explicit prompt structuring, planning-time allocation, and modelling of target syntactic patterns (Pham & Nguyen, 2023). Writing anxiety interventions have employed expressive writing (Klein & Boals, 2001), portfolio-based assessment (Dörnyei & Ryan, 2015), and self-efficacy enhancement protocols (Bandura, 1997). Hiver et al. (2021) argued for dynamic, systems-oriented approaches addressing the co-evolutionary relationship between CL and anxiety. Atasoy and Çepni (2024) demonstrated in a TaFL context that a six-week affective literacy programme reduced FLWA and improved written output; however, that study lacked rigorous experimental controls and did not measure SC explicitly. The present investigation builds on this foundation while addressing these limitations through a larger sample, a longer and more theoretically comprehensive intervention, and a validated SC measurement battery.

Although the preceding sections review Cognitive Load Theory, FLWA research, and CL/anxiety intervention studies as largely parallel literatures, the theoretical rationale for CAWEP rests on a more specific claim about how these constructs jointly constrain SC. Following the cognitive-affective interference model (Eysenck & Derakshan, 2011), FLWA is treated here not as a construct independent of CL but as a source of extraneous load in its own right: intrusive, evaluation-focused cognition consumes working-memory resources that would otherwise be available for syntactic planning, so that extraneous CL and FLWA function as overlapping, mutually reinforcing drains on a common resource pool rather than as separate predictors of SC (cf. Cheng et al., 2026). Under this account, SC is suppressed through two convergent routes: a resource-competition route, in which elevated extraneous load directly limits the working-memory capacity available for constructing embedded and subordinate structures, and an avoidance route, in which anxious writers strategically simplify syntax to reduce the perceived risk of error (Woodrow, 2011; Zhan, 2024). CAWEP was designed to interrupt both routes simultaneously rather than to test them as independent, additive components: the extraneous-load-reduction sessions target

the resource-competition route by removing task-design sources of load unrelated to the syntactic learning goal itself; the FLWA-regulation workshops target the avoidance route by lowering the affective cost associated with syntactic risk-taking; and the germane-load-promotion sessions are intended to convert the working-memory capacity freed by the first two components into durable syntactic schemas, consistent with the element-interactivity principle (van Merriënboer & Sweller, 2022). This account also clarifies why germane-load activities and explicit syntactic instruction proved difficult to separate in the present design: germane load is, by definition, the cognitive investment directed toward schema construction, so any activity that promotes schema-consistent processing of Turkish clause-embedding structures simultaneously functions as syntactic instruction. We did not attempt to isolate germane-load promotion from explicit instruction as independent treatment arms; the implications of this design decision for attributing observed SC gains to a single intervention component are taken up in the Discussion and Limitations.

[Figure 1]

Figure 1. *Conceptual model of the relationships among extraneous cognitive load, foreign language writing anxiety, germane cognitive load, and syntactic complexity, illustrating the pathway through which CAWEP is hypothesised to produce gains in syntactic complexity.*

Figure 1 summarises this working model, depicting extraneous CL and FLWA as overlapping, proximal suppressors of SC and germane CL as the pathway through which CAWEP is hypothesised to produce SC gains.

## 2.4. The TaFL Writing Context

Turkish presents a distinctive typological profile SOV word order, extensive suffixation, vowel harmony, differential object marking, and rich evidential morphology—that collectively imposes high intrinsic CL on learners from non-Altaic L1 backgrounds (Johanson & Csató, 1998; Yıldız, 2022). Producing syntactically complex written Turkish requires not only lexical and morphological automaticity but also schema-based processing of clause-embedding structures that may have no direct L1 equivalent. Against this backdrop, extraneous load generated by ambiguous writing tasks, and anxiety generated by evaluative pressure, are likely to have particularly disruptive effects precisely the conditions targeted by CAWEP.

# 3. Research Questions

The study was guided by the following research questions:

Because pretest group equivalence is a methodological precondition for the between-groups comparisons that follow, rather than a substantive question in its own right, it is reported below as a preliminary equivalence check (RQ1) that precedes the three primary research questions (RQ2–RQ4).

RQ1. Are there statistically significant differences between the experimental and control groups in pretest scores on cognitive load, writing anxiety, and syntactic complexity, confirming group equivalence

prior to intervention?

RQ2. Does participation in the 14-week CAWEP produce significantly lower extraneous cognitive load, lower writing anxiety, and higher syntactic complexity at post-test compared to the control condition, after controlling for pretest scores?

RQ3a. Does FLWA mediate the relationship between extraneous CL and SC at post-test in both groups?

RQ3b. Does CAWEP reduce extraneous CL as a mechanism through which the intervention improves SC?

RQ4. How do experimental group participants qualitatively describe changes in their affective and cognitive processes during syntactically complex writing tasks following CAWEP participation?

# 4. Method

## 4.1. Research Design

A quasi-experimental pretest–posttest non-equivalent control group design (Campbell & Stanley, 1963; Shadish et al., 2002) was employed, appropriate when random assignment of participants to conditions is not feasible, as is typical in intact classroom settings, but when pretest equivalence between groups can be statistically established. The experimental group participated in CAWEP alongside their regular TaFL curriculum; the control group received the same regular curriculum without CAWEP components. The quantitative strand was augmented by a concurrent think-aloud protocol (TA) qualitative strand (Ericsson & Simon, 1993) administered to a purposive 48-participant subsample, yielding a convergent parallel mixed-methods design (Creswell & Creswell, 2023). Integration of strands occurred at the interpretation stage through joint displays (Creamer, 2021).

Internal validity threats were mitigated as follows: selection bias through statistical equivalence testing at pretest; history effects by conducting all sessions within the same academic semester; testing effects through parallel but non-identical writing prompts at pre- and post-test; and instrumentation consistency through standardised administration protocols. Propensity score matching was used as a supplementary check to confirm group equivalence on demographic covariates.

## 4.2. Participants

Participants were recruited from four universities in Türkiye (two in Ankara, one in İstanbul, one in İzmir) that offer accredited TaFL programmes for international students. Eligibility criteria required that participants: (a) were enrolled in a B1–C1-level TaFL course; (b) had completed at least one semester of formal Turkish instruction; (c) were non-heritage learners of Turkish; and (d) provided written informed consent. Intact classes were assigned to conditions: classes at two universities were allocated to the experimental condition, and classes at the other two to the control condition. A total of 300 participants met eligibility criteria and consented (experimental: n = 150; control: n = 150). A purposive subsample of

48 was selected for the TA strand, stratified by condition (24 per group) and proficiency level (8 per level: B1, B2, C1). Demographic information for the full quantitative sample is presented in Table 1.

Table 1
Demographic Characteristics of Experimental and Control Group Participants (N = 300)

| Variable | Category | Experimental (n = 150) | % | Control (n = 150) | % | Total | % |
|---|---|---|---|---|---|---|---|
| Gender | Female | 82 | 54.7 | 80 | 53.3 | 162 | 54.0 |
| | Male | 68 | 45.3 | 70 | 46.7 | 138 | 46.0 |
| Age | 18–22 | 94 | 62.7 | 96 | 64.0 | 190 | 63.3 |
| | 23–27 | 43 | 28.7 | 41 | 27.3 | 84 | 28.0 |
| | 28+ | 13 | 8.7 | 13 | 8.7 | 26 | 8.7 |
| Turkish Proficiency | B1 | 52 | 34.7 | 54 | 36.0 | 106 | 35.3 |
| | B2 | 41 | 27.3 | 41 | 27.3 | 82 | 27.3 |
| | C1 | 21 | 14.0 | 20 | 13.3 | 41 | 13.7 |
| L1 Background | Arabic | 47 | 31.3 | 48 | 32.0 | 95 | 31.7 |
| | Slavic languages | 36 | 24.0 | 36 | 24.0 | 72 | 24.0 |
| | Persian | 27 | 18.0 | 27 | 18.0 | 54 | 18.0 |
| | East Asian | 23 | 15.3 | 22 | 14.7 | 45 | 15.0 |
| | Other | 17 | 11.3 | 17 | 11.3 | 34 | 11.3 |

Note. Turkish proficiency levels were determined by TYT scores and course enrolment level. Chi-square tests confirmed no significant between-group differences (all $p > .05$).

*[Insert* Table 1 *here]*

# 4.3. The Intervention: CAWEP

CAWEP is a 14-week, theoretically grounded intervention delivered by trained TaFL instructors in weekly 90-minute sessions alongside regular curriculum content. The programme was constructed on three mutually reinforcing principles: (1) extraneous CL reduction through explicit task structuring and scaffolding; (2) germane CL promotion through syntactic schema-building and metalinguistic awareness activities; and (3) FLWA regulation through evidence-based affective self-management strategies. Table 2 presents the session-by-session programme overview.

Table 2
Session-by-Session Overview of CAWEP

| Week | Session | Content | Target Component |
|---|---|---|---|
| 1 | Pre-test & Orientation | CLS, FLCAS-W, writing pre-test; protocol introduction | Baseline |
| 2 | Task Design Training I | Structured prompts; worked outline activities | Extraneous CL ↓ |
| 3 | Syntax Schema Building I | T-unit and embedded clause pattern practice | Germane CL ↑ |
| 4 | Affective Regulation I | Anxiety management strategies; writing reframing | Anxiety ↓ |
| 5 | Task Design Training II | Staged writing tasks; peer feedback protocols | Extraneous CL ↓ |
| 6 | Syntax Schema Building II | Subordinate clause types; authentic text analysis | Germane CL ↑ |
| 7 | Affective Regulation II | Metacognitive diary; attribution techniques | Anxiety ↓ |
| 8 | Integrated Practice | Authentic argumentative writing | All components |
| 9 | Formative Assessment | Formative writing + CLS and anxiety mid-test | Mid-test |
| 10 | Task Design III (Advanced) | Multi-dimensional tasks; reduced scaffolding | Extraneous CL ↓ |
| 11 | Syntax Schema III (Advanced) | Academic Turkish structures; conjunction variety | Germane CL ↑ |
| 12 | Affective Regulation III | Self-efficacy enhancement; writing portfolio | Anxiety ↓ |
| 13 | Review & Feedback | Individual writing conferences; SC awareness training | Consolidation |
| 14 | Post-test | CLS, FLCAS-W, writing post-test; exit survey | Post-test |

Note. CLS = Cognitive Load Scale; FLCAS-W = Foreign Language Classroom Anxiety Scale – Writing Subscale. The control group followed the standard TaFL curriculum without CAWEP components.

*[Insert* Table 2 *here]*

Extraneous CL Reduction Sessions (Weeks 2, 5, 10) operationalised CLT's extraneous load reduction principles (Sweller et al., 2019) through: (a) provision of explicitly structured, CEFR-aligned writing prompts with clearly delineated task parameters; (b) segmented task phases with explicit time allocations; (c) worked example pairs demonstrating syntactically simple versus complex responses;

and (d) collaborative planning activities that externalise the cognitive work of text organisation, reducing the burden on WM during drafting.

Germane CL Promotion Sessions (Weeks 3, 6, 11) targeted schema formation through: (a) pattern-based practice in T-unit construction and embedded subordinate clause types; (b) analysis of authentic academic Turkish texts; (c) syntactic paraphrasing exercises requiring transformation of coordination into subordination; and (d) interleaved practice to promote schema generalisation (van Merriënboer & Sweller, 2022).

FLWA Regulation Workshops (Weeks 4, 7, 12) drew on cognitive-behavioural and positive psychology approaches (Dörnyei & Ryan, 2015; Hiver et al., 2021) through: (a) psychoeducation on the CL–anxiety interface; (b) reframing of grammatical error as informative feedback; (c) metacognitive diary writing to identify anxiety triggers; (d) collaborative writing protocols; and (e) portfolio-based progress documentation to build evidential self-efficacy.

All four CAWEP instructors participated in a three-day training workshop prior to the intervention. Fidelity checks via structured classroom observation (two sessions per instructor) confirmed adequate implementation, with a mean fidelity score of 91.4%.

# 4.4. Data Collection Instruments

Cognitive load was measured with the Cognitive Load Scale (CLS) developed by Kalyuga and Sweller (2024) and adapted into Turkish by Arslan and Yıldırım (2024). The scale consists of 21 items across three subscales (Intrinsic, Extraneous, Germane; 7 items each) rated on a 7-point Likert scale. The Turkish adaptation reported Cronbach's α = .87, .84, and .81 for the respective subscales and confirmed the three-factor structure (CFI = .94, RMSEA = .058). In the present study, pretest α values were .86, .84, and .80; post-test values were .85, .83, and .81, indicating measurement stability. Test-retest reliability over two weeks was r = .82.

Writing anxiety was measured with the 15-item FLCAS-W validated for TaFL contexts by Demir and Özkan (2023; n = 412; α = .91; CFI = .96, RMSEA = .053; test-retest r = .84). In the present study, pretest α = .89 and post-test α = .90. Convergent validity was confirmed through significant correlations with the FLCAS total score (r = .74, p < .001).

Syntactic complexity was operationalised via three indices derived from argumentative writing samples: (a) an SC composite (mean of z-standardised MLT and embedded clause rate; Lu, 2017); (b) mean length of T-unit (MLT); and (c) embedded clause rate (EmbRate). T-unit segmentation followed protocols established by Hunt (1965) and adapted for TaFL by Güneş and Yıldız (2024). The TAASSC tool (Kyle, 2021) was used for automated SC extraction and manually verified. Inter-rater reliability was κ = .88 (pre-test) and κ = .89 (post-test).

The think-aloud protocol (TA) instrument consisted of a concurrent think-aloud administered to the 48-participant subsample during a 45-minute argumentative writing task. Training followed Ericsson and

Simon (1993). Transcripts were analysed using thematic analysis (Braun & Clarke, 2022) with a deductive-dominant approach. The protocol was administered individually in a quiet testing room by a trained research assistant who first modelled the technique using a brief non-target warm-up task, then instructed participants to verbalise their thoughts continuously while completing Prompt C; sessions were audio-recorded and transcribed verbatim. Integration with the quantitative strand occurred at the design stage, through purposive stratification of the TA subsample by condition and proficiency level, and at the interpretation stage, through joint displays juxtaposing thematic codes with the condition-level ANCOVA and mediation results reported in Sections 5.2 and 5.3. Member checking with 10 participants and inter-coder reliability (κ = .83) ensured credibility.

## 4.5. Data Collection Procedure

Ethical approval was obtained from the Scientific Research Ethics Committee of Ağrı İbrahim Çeçen University (Protocol No. 2024-EK-0187). Data collection unfolded across four stages. Stage 1 (Week 1): all 300 participants completed the CLS, FLCAS-W, and a 40-minute argumentative writing task (Prompt A). Stage 2 (Weeks 2–13): the experimental group participated in CAWEP; a mid-test (CLS + FLCAS-W only) was administered at Week 9. The control group followed the standard TaFL curriculum. Stage 3 (Week 14): all participants completed the CLS, FLCAS-W, and a 40-minute writing task (Prompt B); TA sessions were conducted with the 48-participant subsample using Prompt C. Stage 4 (Weeks 15–16): post-test writing samples were processed through TAASSC and manually verified.

## 4.6. Data Analysis

Quantitative analyses were conducted in SPSS 28.0 and R 4.3.2. Data screening confirmed normal distributions (Shapiro–Wilk), absence of multivariate outliers (Mahalanobis distance), and minimal missing data (< 1.5%, resolved via multiple imputation). Pretest group equivalence was assessed through independent-samples t-tests and Cohen's d (RQ1). Primary intervention effects were tested using ANCOVA with post-test scores as the dependent variable, group as the between-subjects factor, and corresponding pretest scores as covariates (RQ2). Partial eta-squared ($\eta^2p$) was the effect size index, with .01, .06, and .14 interpreted as small, medium, and large (Cohen, 1988). Bootstrapped mediation analyses (10,000 resamples; R mediation package) tested the mediational role of FLWA in the CL–SC relationship (RQ3). Qualitative TA transcripts were analysed through thematic analysis in NVivo 14 (RQ4).

# 5. Results

## 5.1. RQ1: Pretest Group Equivalence

Independent-samples t-tests confirmed statistical equivalence on all dependent variables at baseline (Table 3). CL total scores (Experimental: M = 78.21, SD = 14.43; Control: M = 78.96, SD = 13.87; t = − 0.47, p = .638, d = .06), writing anxiety scores (Experimental: M = 82.47, SD = 16.91; Control: M = 83.12, SD = 16.55; t = − 0.34, p = .731, d = .04), and SC composite scores (Experimental: M = 9.71, SD = 2.48; Control:

M = 9.84, SD = 2.61; t = − 0.44, p = .660, d = .05) showed no significant differences. All Cohen's d values fell below .10, confirming negligible between-group differences and satisfying the fundamental internal validity requirement.

Table 3
Pretest Descriptive Statistics and t-Test Results (N = 300)

| Variable | Exp. M | Exp. SD | Ctrl. M | Ctrl. SD | t | p | d |
|---|---|---|---|---|---|---|---|
| CLS Total | 78.21 | 14.43 | 78.96 | 13.87 | −0.47 | .638 | .06 |
| Intrinsic Load | 26.04 | 5.91 | 26.31 | 5.72 | −0.41 | .682 | .05 |
| Extraneous Load | 28.87 | 6.22 | 29.14 | 6.09 | −0.38 | .704 | .04 |
| Germane Load | 23.30 | 5.51 | 23.51 | 5.43 | −0.34 | .734 | .04 |
| FLCAS-W | 82.47 | 16.91 | 83.12 | 16.55 | −0.34 | .731 | .04 |
| SC Composite | 9.71 | 2.48 | 9.84 | 2.61 | −0.44 | .660 | .05 |
| MLT | 8.52 | 2.09 | 8.67 | 2.17 | −0.62 | .535 | .07 |
| Embedded Clause Rate | 0.41 | 0.17 | 0.42 | 0.18 | −0.50 | .618 | .06 |

Note. CLS = Cognitive Load Scale; FLCAS-W = Foreign Language Classroom Anxiety Scale – Writing Subscale; SC = Syntactic Complexity; MLT = Mean Length of T-unit. Significance threshold adjusted to $p < .006$ (Bonferroni correction). All d values are negligible ($|d| < .10$).

*[Insert* Table 3 *here]*

# 5.2. RQ2: Post-test Between-Group Differences

Post-test comparisons (Table 4) revealed that CAWEP produced large and statistically significant advantages for the experimental group across all targeted outcome variables. The experimental group obtained markedly lower extraneous CL scores (Experimental: M = 20.34, SD = 4.62; Control: M = 28.61, SD = 5.87; F = 108.47, $p < .001$, $\eta^2p = .27$), reported significantly lower writing anxiety (Experimental: M = 61.23, SD = 13.45; Control: M = 81.87, SD = 15.63; F = 97.81, $p < .001$, $\eta^2p = .25$), and demonstrated considerably higher SC composite scores (Experimental: M = 14.38, SD = 2.71; Control: M = 10.12, SD = 2.43; F = 129.44, $p < .001$, $\eta^2p = .30$).

Table 4
Post-test Descriptive Statistics and ANCOVA Results (N = 300)

| Variable | Exp. M | Exp. SD | Ctrl. M | Ctrl. SD | F | p | $\eta^2p$ |
|---|---|---|---|---|---|---|---|
| CLS Total | 62.18 | 11.74 | 77.43 | 13.21 | 98.34 | < .001 | .25 |
| Intrinsic Load | 24.91 | 4.83 | 25.87 | 5.14 | 2.87 | .091 | .01 |
| Extraneous Load | 20.34 | 4.62 | 28.61 | 5.87 | 108.47 | < .001 | .27 |
| Germane Load | 16.93 | 3.88 | 22.95 | 4.71 | 84.62 | < .001 | .22 |
| FLCAS-W | 61.23 | 13.45 | 81.87 | 15.63 | 97.81 | < .001 | .25 |
| SC Composite | 14.38 | 2.71 | 10.12 | 2.43 | 129.44 | < .001 | .30 |
| MLT | 12.41 | 2.34 | 8.93 | 2.08 | 114.67 | < .001 | .28 |
| Embedded Clause Rate | 0.64 | 0.19 | 0.43 | 0.16 | 67.31 | < .001 | .18 |

Note. ANCOVA controls for pretest scores as covariate. F values based on adjusted means. $\eta^2p \geq .14$ = large effect (Cohen, 1988).

*[Insert* Table 4 *here]*

These effect sizes substantially exceed the medium-to-large thresholds reported in prior FL writing intervention meta-analyses (Huang, 2018; Zhan, 2024). Within-group change analyses (Table 5) confirm the asymmetric pattern: experimental group improvements were large and significant (all $p < .001$), while control group changes were non-significant and negligible (all $p > .18$).

Table 5
Within-Group Pre-to-Post Change Statistics

| Variable | Group | Pre-test M | Post-test M | ΔM | t | p |
|---|---|---|---|---|---|---|
| CLS Total | Experimental | 78.21 | 62.18 | −16.03 | 18.74 | < .001 |
| | Control | 78.96 | 77.43 | −1.53 | 1.12 | .264 |
| FLCAS-W | Experimental | 82.47 | 61.23 | −21.24 | 20.43 | < .001 |
| | Control | 83.12 | 81.87 | −1.25 | 0.91 | .365 |
| SC Composite | Experimental | 9.71 | 14.38 | + 4.67 | 22.81 | < .001 |
| | Control | 9.84 | 10.12 | + 0.28 | 1.33 | .185 |
| MLT | Experimental | 8.52 | 12.41 | + 3.89 | 21.64 | < .001 |
| | Control | 8.67 | 8.93 | + 0.26 | 1.18 | .240 |
| Emb. Clause Rate | Experimental | 0.41 | 0.64 | + 0.23 | 16.93 | < .001 |
| | Control | 0.42 | 0.43 | + 0.01 | 0.44 | .661 |

Note. ΔM = post-test minus pretest mean difference.

*[Insert* Table 5 *here]*

# 5.3. RQ3: Mediation Analysis

Bootstrapped mediation analyses (Table 6) confirmed that writing anxiety partially mediates the relationship between extraneous CL and SC in both conditions. The indirect effect was somewhat smaller in the experimental group (b = − .173, 95% CI [− .254, − .097], p = .002) than in the control group (b = − .198, 95% CI [− .289, − .112], p = .004), indicating that CAWEP attenuated anxiety-based suppression of SC without fully eliminating the CL–anxiety–SC pathway. The third path demonstrated that intervention group membership reduced extraneous CL, which in turn predicted SC gains (b = .241, 95% CI [.151, .335], p = .001). All indirect effects are significant as confidence intervals exclude zero.

Table 6
Bootstrapped Indirect Effects (N = 300, 10,000 Resamples)

| Path | b | Boot SE | 95% CI | p | Mediation |
|---|---|---|---|---|---|
| Extraneous CL → Anxiety → SC (Experimental) | −.173 | .041 | [− .254, − .097] | .002 | Partial |
| Extraneous CL → Anxiety → SC (Control) | −.198 | .046 | [− .289, − .112] | .004 | Partial |
| Group → Extraneous CL → SC | .241 | .048 | [.151, .335] | .001 | Partial |

Note. b = unstandardised indirect effect; Boot SE = bootstrapped standard error; CI = 95% bias-corrected confidence interval.

*[Insert* Table 6 *here]*

# 5.4. RQ4: Qualitative Findings

Thematic analysis of post-test TA transcripts from the 48-participant subsample yielded three superordinate themes in the experimental group largely absent or attenuated in control group transcripts as summarised in Table 7.

Table 7
Qualitative Themes and Representative Excerpts from Think-Aloud Protocols (Experimental Group, n = 24)

| Theme | Sub-Theme | Representative Excerpt | Mediation Path |
|---|---|---|---|
| CL Reduction | Task clarity | 'I can plan what to write more easily now; there is no confusion.' | Extraneous CL ↓ → SC ↑ |
| CL Reduction | Structured planning | 'When I outline first, complex sentences come naturally.' | Germane CL ↑ → SC ↑ |
| Anxiety Regulation | Fear of evaluation | 'I no longer worry about making mistakes; I just write.' | Anxiety ↓ → SC ↑ |
| Anxiety Regulation | Syntactic avoidance | 'I am trying complex sentences because I am less afraid now.' | Anxiety ↓ → SC ↑ |
| Germane Strategy | Metalinguistic awareness | 'I am consciously thinking about where to place the dependent clause.' | Germane CL ↑ → SC ↑ |

Note. All excerpts from experimental group participants across B1–C1 proficiency levels.

*[Insert* Table 7 *here]*

Experimental group participants consistently described how the clarity of task instructions facilitated cognitive planning and how syntactic avoidance had been replaced by deliberate syntactic risk-taking. These themes were markedly less prevalent in the control group TA data. The exclusive emergence of the germane strategy theme characterised by conscious metalinguistic monitoring in experimental

group transcripts indicates that CAWEP generated not only load and anxiety reductions but also enhanced metacognitive self-regulation during writing.

# 6. Discussion

This study was designed, implemented, and evaluated CAWEP, a 14-week, theoretically grounded intervention targeting extraneous CL, FLWA, and SC in TaFL writing among 300 international students. The findings provide strong and convergent evidence supporting the programme's efficacy and illuminate the mechanisms through which intervention effects operate.

ANCOVA results demonstrated large, significant post-test advantages for the experimental group across all outcome variables. These effect sizes substantially exceed thresholds reported in prior FL writing intervention meta-analyses (Huang, 2018; Zhan, 2024), suggesting that the multi-component, theory-driven nature of CAWEP amplifies its effectiveness relative to single-component interventions. The programme's coherent integration of CL reduction, germane load promotion, and anxiety regulation appears to create a mutually reinforcing system in which progress on one dimension facilitates progress on others.

These effect sizes, however, warrant a more critical reading than a straightforward efficacy interpretation permits. Magnitudes in the $\eta^2p$ = .25–.30 range are unusually large for a 14-week classroom intervention and exceed not only prior FL writing-anxiety meta-analyses (Huang, 2018; Zhan, 2024) but also the effect sizes typically reported for single-component CLT interventions in other domains (Sweller et al., 2019). Three non-mutually-exclusive alternative explanations deserve consideration alongside a genuine treatment effect. First, because entire universities — not individuals — were assigned to condition, any pre-existing but unmeasured institutional difference in student motivation, prior curriculum quality, or assessment norms is perfectly confounded with condition and could inflate the apparent effect of CAWEP. Second, the experimental group's awareness of receiving a novel, resource-intensive programme delivered by specially trained instructors may have generated demand characteristics or a Hawthorne-type response independent of the programme's specific theoretical content. Third, the self-reported CL and FLWA measures are susceptible to social-desirability pressure that could be differentially elevated in a condition explicitly framed around load and anxiety reduction. None of these alternatives can be ruled out by the present non-equivalent control group design; we therefore treat the reported effect sizes as an upper-bound estimate of CAWEP's likely impact under closely monitored implementation conditions rather than as a definitive index of its effectiveness in ordinary practice.

Mediation analyses confirmed that writing anxiety partially mediates the extraneous CL–SC relationship in both conditions, replicating and extending correlational findings of prior research (Woodrow, 2011; Zhan, 2024) in an experimental context. The smaller indirect effect in the experimental group (b = − .173 versus b = − .198 in controls) indicates that CAWEP attenuated anxiety-based suppression of SC—not by eliminating the CL–anxiety–SC pathway, but by reducing the magnitude of anxiety activation for any given level of extraneous load. This is consistent with the cognitive-affective interference model

(Eysenck & Derakshan, 2011; Hiver et al., 2021): when anxiety regulation capacity is strengthened through intervention, the disruptive impact of residual extraneous load is buffered.

This mediational account is theoretically coherent, but it should not be read as establishing the causal ordering it implies. The model treats extraneous CL as prior to FLWA, which is in turn treated as prior to SC; however, all three variables were measured concurrently at post-test, and an equally defensible alternative — consistent with the cognitive-affective interference literature itself (Eysenck & Derakshan, 2011) — would specify FLWA as antecedent to, rather than a consequence of, extraneous CL, since anxious anticipation of a writing task can itself generate extraneous load before task engagement begins. The bootstrapped confidence intervals in Table 6 index the statistical robustness of the indirect path under the assumed causal ordering; they cannot adjudicate between competing causal orderings, and the smaller indirect effect in the experimental relative to the control group is consistent with several distinct causal stories, not only the one advanced here. Cross-lagged or multi-wave panel designs would be required to test these alternatives directly, and we flag this as a priority for future replications of CAWEP.

The TA protocol data provided granular process-level evidence corroborating these structural paths. The emergence of syntactic risk-taking behaviour exclusively among experimental participants, and its absence from control TA data, suggests that CAWEP produced a qualitative shift in learners' orientation toward syntactic complexity from a domain of avoidance to one of strategic engagement. The metalinguistic awareness theme indicates that CAWEP's germane load activities promoted not only schema formation but also metacognitive monitoring of syntactic production.

The unexpectedly lower germane load scores in the experimental group at post-test align with CLT's learning efficiency model (Kalyuga & Sweller, 2024): as syntactic schemas become automated, effortful schema building demands less conscious allocation, resulting in lower self-reported germane load even as written output quality improves a validity indicator observed in CLT-based intervention studies in other domains (Leow, 2023).

Several limitations warrant acknowledgement. Non-random assignment of classes introduces selection confounds not fully addressable through statistical equivalence testing or propensity score matching. The intervention was delivered by four instructors across four institutions, introducing instructor and institutional variation despite fidelity monitoring; future studies should explicitly model instructor as a random factor. In the present study, each experimental class was taught by a single, non-rotated instructor, so instructor effects cannot be statistically separated from either institutional or treatment effects; the effect sizes reported here should accordingly be interpreted as attributable to the CAWEP-instructor-institution package as implemented, rather than to the CAWEP curriculum in isolation. The single-semester timeframe limits conclusions about the durability of CAWEP effects. Finally, the exclusive use of argumentative writing limits generalisability to other TaFL writing genres.

# 7. Conclusion

This study furnishes the first experimental evidence for the efficacy of a cognitively and affectively integrated writing intervention in the TaFL context. By simultaneously targeting extraneous CL, writing anxiety, and germane CL through CAWEP, large and significant improvements in SC were produced among 300 international students relative to a business-as-usual control condition. Mediation analyses revealed that effects on SC operated both directly, through extraneous load reduction, and indirectly, through attenuation of the anxiety-mediated suppression of SC. Concurrent think-aloud protocol data provided convergent process-level evidence, revealing qualitative shifts in learners' syntactic engagement and metacognitive awareness.

These findings advance the field in three respects. Theoretically, they establish that CLT and FLWA operate as structurally interconnected rather than independent determinants of SC in TaFL writing, and that both are tractable through targeted pedagogical intervention. Methodologically, they demonstrate the added value of integrating pretest–posttest experimental designs with bootstrapped mediation analysis and concurrent TA protocols in FL intervention research. Practically, they provide TaFL practitioners and curriculum designers with an evidence-based, replicable programme specification implementable within existing institutional constraints.

## 8. Pedagogical Recommendations

TaFL writing instructors should consistently provide explicit, well-structured prompts with clear parameters, worked syntactic examples, and segmented planning–drafting–revision phases. Given the typological distance between Turkish and most learner L1s, task formats that externalise organisational demands are particularly warranted (Sweller et al., 2019). Portfolio-based assessment, collaborative writing, and metacognitive diary protocols should be treated not as supplementary activities but as core curricular components with dedicated instructional time. Interleaved practice across Turkish clause types, combined with syntactic paraphrasing tasks and authentic text analysis, provides the schema-generalisation experiences necessary for durable syntactic automaticity; these activities should be sequenced after extraneous load has been reduced (van Merriënboer & Sweller, 2022). Future research should pursue longitudinal follow-up studies, randomised controlled trials, multimodal CL measurement, and adaptive versions of CAWEP across other TaFL writing genres and online instructional contexts.

## Declarations

# Ethics Approval and Consent to Participate

All procedures were conducted in full accordance with the ethical standards of the Scientific Research Ethics Committee of Ağrı İbrahim Çeçen University and the 1964 Declaration of Helsinki and its subsequent revisions. Written informed consent was obtained from all individual participants prior to data collection. Participants were informed of their right to withdraw at any time without consequence.

**Consent to Publish**

All participants provided written consent for their anonymised data to be used for research purposes and published in academic outlets.

# Competing Interests

The authors declare that they have no competing interests.

# Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

# Author Contribution

K.K. → Dr. Kılıç KÖÇERİA.A. → Prof. Dr. Ahmet AKÇAYA.A. → Adnan AYDINTÜRKA.A. → Akif Arslan

# Acknowledgements

The authors wish to thank the participating institutions and their international student cohorts for their contributions to this research.

# Data Availability

The datasets supporting the conclusions of this article are available from the corresponding author upon reasonable request.

## References

1. Arslan H, Yıldırım S. Bilişsel Yük Ölçeği'nin Türkçe uyarlaması [Turkish adaptation of the Cognitive Load Scale]. Eğitim ve Bilim. 2024;49(218):1–22. https://doi.org/10.15390/EB.2024.13547.
2. Atasoy R, Çepni S. Affective literacy intervention in TaFL writing: A six-week programme evaluation. Dil ve Dilbilim Araştırmaları. 2024;20(2):143–69. https://doi.org/10.17263/jlls.2024.2.07.
3. Ateş A, Bahşi N. The effects of emotional intelligence of foreign students learning Turkish on speaking and writing anxiety. Educ Q Reviews. 2022;5(4):112–26. https://doi.org/10.31014/aior.1993.05.04.582.
4. Bandura A. Self-efficacy: The exercise of control. W. H. Freeman; 1997.
5. Batmaz Ö, Büyüikikiz KK. Yabancı dil olarak Türkçe öğretiminde duyusal faktörlerin anlama becerileri içindeki yeri. Aydın Tömer Dil Dergisi. 2025;10(2):433–69.

https://doi.org/10.17932/IAU.TOMER.2016.019/tomer_v010i2008.

6. Bensoussan M, Kreindler I. Improving advanced reading comprehension in a foreign language. J Res Reading. 1990;13(1):55–68. https://doi.org/10.1111/j.1467-9817.1990.tb00340.x.
7. Braun V, Clarke V. Thematic analysis: A practical guide. SAGE; 2022.
8. Campbell DT, Stanley JC. Experimental and quasi-experimental designs for research. Houghton Mifflin; 1963.
9. Cheng S, Lin C, Wei D, Chen Y. A longitudinal psychological model of writing development: Interactions among self-efficacy, regulation, and anxiety. Front Psychol. 2026;17:1757045. https://doi.org/10.3389/fpsyg.2026.1757045.
10. Cheng Y-S. A measure of second language writing anxiety: Scale development and preliminary validation. J Second Lang Writ. 2004;13(4):313–35. https://doi.org/10.1016/j.jslw.2004.07.001.
11. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Lawrence Erlbaum; 1988.
12. Creamer EG. Advancing grounded theory with mixed methods. Routledge; 2021.
13. Creswell JW, Creswell JD. Research design: Qualitative, quantitative, and mixed methods approaches. 6th ed. SAGE; 2023.
14. Daly JA, Miller MD. The empirical development of an instrument to measure writing apprehension. Res Teach Engl. 1975;9(3):242–9.
15. Demir Y, Özkan Y. Yabancı dil yazma kaygısı ölçeği: TaFL bağlamına uyarlanması [Adaptation to the TaFL context]. Dil Dergisi. 2023;174(2):45–68. https://doi.org/10.33690/dilder.2023.14.
16. Dörnyei Z, Ryan S. The psychology of the language learner revisited. Routledge; 2015.
17. Ericsson KA, Simon HA, Rev. MIT Press.
18. Eysenck MW, Derakshan N. New perspectives in attentional control theory. Pers Indiv Differ. 2011;50(7):955–60. https://doi.org/10.1016/j.paid.2010.08.019.
19. Güneş F. Yabancı dil olarak Türkçe öğretiminde sözdizimsel gelişim. Türkçe Öğretimi Araştırmaları Dergisi. 2021;7(1):14–33. https://doi.org/10.29065/toad.2021.7.1.02.
20. Güneş F, Yıldız C. T-birim tabanlı sözdizimsel karmaşıklık normları. Dil ve Dilbilim Araştırmaları. 2024;20(1):88–112. https://doi.org/10.17263/jlls.2024.1.04.
21. Hiver P, Al-Hoorie AH, Larsen-Freeman D. (2021). Dynamic systems theory and L2 pedagogy. In Language learning and teaching as complex and dynamic systems (pp. 1–25). Multilingual Matters. https://doi.org/10.21832/9781800410466-003
22. Horwitz EK, Horwitz MB, Cope J. Foreign language classroom anxiety. Mod Lang J. 1986;70(2):125–32. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x.
23. Huang J. A meta-analysis of the effect of writing anxiety on English writing quality. J Second Lang Writ. 2018;39:87–96. https://doi.org/10.1016/j.jslw.2018.01.003.
24. Hunt KW. Grammatical structures written at three grade levels. National Council of Teachers of English; 1965.
25. Johanson L, Csató ÉÁ, editors. The Turkic languages. Routledge; 1998.

26. Kalyuga S, Sweller J. Revising the cognitive load scale: Theoretical refinements and empirical support. Educational Psychol Rev. 2024;36(1). Article 12. https://doi.org/10.1007/s10648-024-09843-0.
27. Klein K, Boals A. Expressive writing can increase working memory capacity. J Exp Psychol Gen. 2001;130(3):520–33. https://doi.org/10.1037/0096-3445.130.3.520.
28. Kormos J. The role of individual differences in L2 writing. J Second Lang Writ. 2012;21(4):390–403. https://doi.org/10.1016/j.jslw.2012.09.003.
29. Kyle K. TAASSC: A tool for the automatic analysis of syntactic sophistication and complexity. Behav Res Methods. 2021;53(4):1735–54. https://doi.org/10.3758/s13428-020-01542-y.
30. Leow RP. Cognitive load theory and SLA: Relevance, limitations, and future directions. Lang Teach Res. 2023;27(3):612–30. https://doi.org/10.1177/13621688211046928.
31. Lin C-C. Impact of visual pedagogical interventions on EFL writing: A mixed-methods analysis of anxiety mitigation and advancement of creative thinking skills. Stud Educational Evaluation. 2026;89. https://doi.org/10.1016/j.stueduc.2026.101619. Article 101619.
32. Lu X. Automated measurement of syntactic complexity in corpus-based L2 writing research. Lang Test. 2017;34(4):493–511. https://doi.org/10.1177/0265532217710675.
33. Pham TH, Nguyen LT. (2023). Scaffolding effects on cognitive load and syntactic complexity in EFL writing. System, 112, Article 102976. https://doi.org/10.1016/j.system.2023.102976
34. Porte G, Richards JC. Replication research in applied linguistics. 2nd ed. Cambridge University Press; 2022.
35. Révész A, Michel M, Lee M. Investigating the relationship between L2 writing processes, L2 proficiency, and writing task complexity. TESOL Q. 2017;51(2):408–21. https://doi.org/10.1002/tesq.294.
36. Shadish WR, Cook TD, Campbell DT. Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin; 2002.
37. Sweller J. Cognitive load during problem solving: Effects on learning. Cogn Sci. 1988;12(2):257–85. https://doi.org/10.1207/s15516709cog1202_4.
38. Sweller J, van Merriënboer JJG, Paas F. Cognitive architecture and instructional design: 20 years later. Educational Psychol Rev. 2019;31(2):261–92. https://doi.org/10.1007/s10648-019-09465-5.
39. van Merriënboer JJG, Sweller J. Cognitive load theory in health professional education: Design principles and strategies. Med Educ. 2022;44(1):85–93. https://doi.org/10.1111/j.1365-2923.2009.03498.x.
40. Woodrow L. College English writing affect: Self-efficacy and anxiety. System. 2011;39(4):510–22. https://doi.org/10.1016/j.system.2011.10.017.
41. Yıldız C. Türkçenin yabancı dil olarak öğretiminde biçim-sözdizim arayüzü. Türk Dili Araştırmaları Yıllığı – Belleten. 2022;70(1):201–28. https://doi.org/10.32010/buluten.2022.19.

42. Yıldız C, Yıldız M. Agglutination and cognitive load in TaFL learners. Lang Learn. 2023;73(2):415–52. https://doi.org/10.1111/lang.12541.
43. Zhan J. Foreign language writing anxiety and writing quality: An updated meta-analysis. Lang Teach Res. 2024;28(1):3–29. https://doi.org/10.1177/13621688211065901.

# Figures

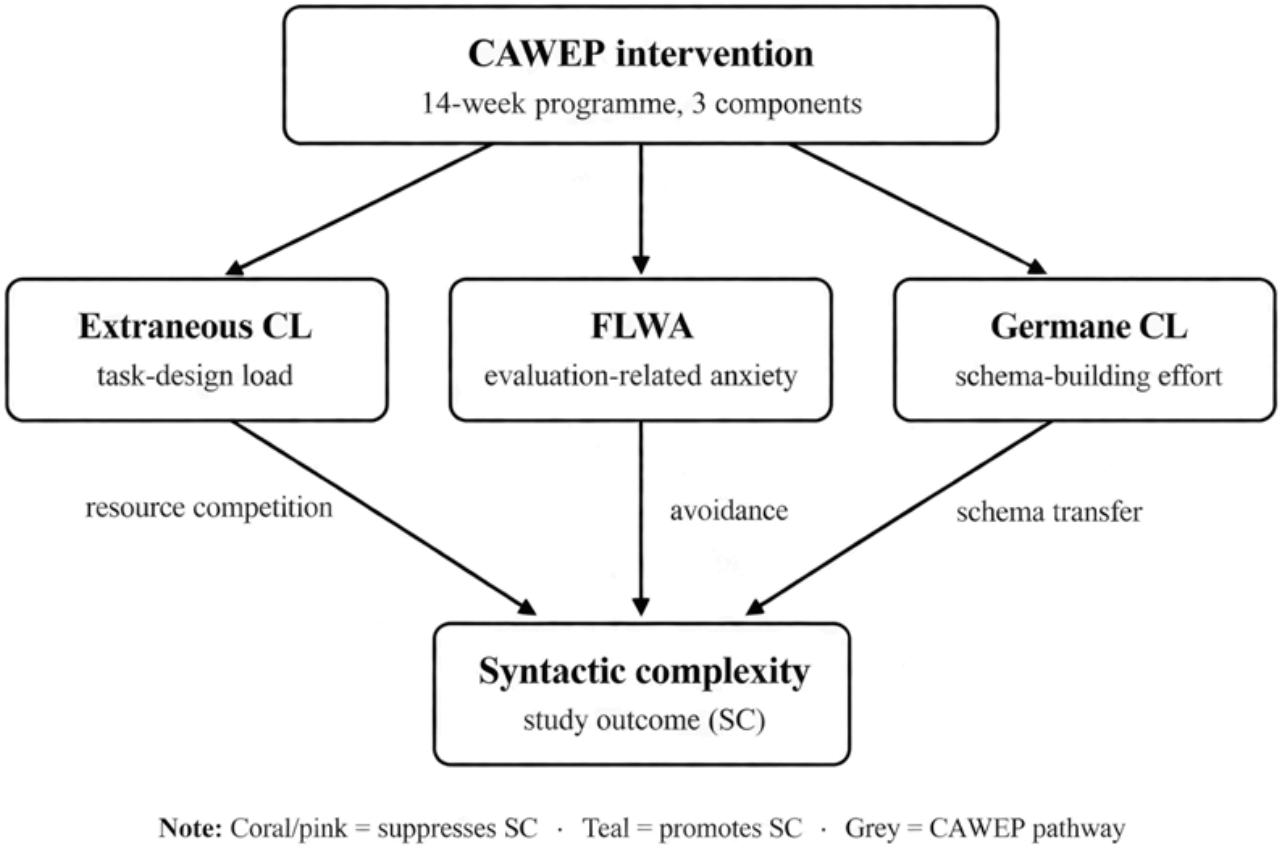



**Figure 1**

Conceptual framework linking CAWEP, cognitive load, writing anxiety, and syntactic complexity

*Conceptual model of the relationships among extraneous cognitive load, foreign language writing anxiety, germane cognitive load, and syntactic complexity, illustrating the pathway through which CAWEP is hypothesised to produce gains in syntactic complexity.*