

# Metacognitive Support Needs in Digital Learning Environments for Critical Writing among Chinese Undergraduates

Jinjin Jiang<sup>1</sup>, Parama Kwangmuang<sup>1</sup>

<sup>1</sup>Department of Innovation, Technology and Learning Science, Faculty of Education, Khon Kaen University, Khon Kaen Province, Thailand

Correspondence: Parama Kwangmuang, Department of Innovation, Technology and Learning Science, Faculty of Education, Khon Kaen University, Khon Kaen Province, Thailand ORCID: <https://orcid.org/0000-0001-9623-0609>  
E-mail: [parakw@kku.ac.th](mailto:parakw@kku.ac.th)

Received: August 17, 2026

Accepted: October 10, 2026

Online Published: October 13, 2026

doi:10.11114/jets.v15i1.9200

URL: <https://doi.org/10.11114/jets.v15i1.9200>

## Abstract

Digital learning environments in Chinese higher education deliver content but often provide limited support for regulating writing. This survey of 400 Chinese undergraduates examined which metacognitive supports in critical writing were perceived as insufficient and whether needs varied across the sample. Exploratory factor analysis, paired discrepancy analysis, latent profile analysis, and two one-sided equivalence tests showed three dimensions: perceived provision, support need, and critical writing self-efficacy. Support need exceeded provision by 0.38 scale points, and 65.5% of students reported unmet need. Structured guidance, progress monitoring, and planning tools were the highest priorities. Needs were equivalent across institution, gender, and prior online learning experience at  $d = \pm 0.40$ . The findings support a common scaffolding architecture with optional intensification for individual learners.

**Keywords:** metacognitive regulation, digital learning environment, critical writing, needs assessment, equivalence testing, Chinese higher education

## 1. Introduction

Chinese undergraduates in critical writing courses use digital learning environments that supply content but give little help with regulating writing. A platform can host readings, collect submissions, and display completion rates without helping students set rhetorical goals, detect weak evidence–claim links, or judge drafts against tacit disciplinary criteria (Pintrich, 2000; Winne & Hadwin, 1998; Zimmerman, 2002) and knowledge-transforming composition (Bereiter & Scardamalia, 1987; Flower & Hayes, 1981). The problem is not absent digital infrastructure but a mismatch between available tools and the metacognitive work critical writing demands.

International research shows that computer-based environments can scaffold self-regulation when they offer planning prompts, monitoring cues, feedback, and reflection support, but these features do not automatically become metacognitive support unless learners can interpret them and act on them.

Three limitations define the gap. First, analytics dashboards often make activity visible without grounding displays in self-regulated learning. Second, automated feedback is useful but unevenly taken up by learners. Third, Chinese higher-education studies have examined online learning and writing strategies, but few digital-environment models organise support around the regulatory cycle for critical writing.

The overarching question is: what metacognitive support do Chinese undergraduates in critical writing courses perceive as insufficient in their digital learning environment, and are those perceptions uniform across the student population? Because the design is cross-sectional, with no baseline or comparison condition, the study is exploratory and advances descriptive sub-questions:

**RQ1.** What dimensional structure underlies Chinese undergraduate students' perceptions of their digital learning environment and their metacognitive support needs in critical writing?

**RQ2.** To what extent does students' perceived need for metacognitive learning support exceed the support currently provided by their digital learning environment, and which specific forms of support are prioritised?

**RQ3.** To what extent do these perceived needs vary across institution, gender, and prior online learning experience?

The study contributes by quantifying the need-provision gap, ranking priority supports, showing that competence beliefs separate from environmental demand, and testing whether a single scaffolding architecture is defensible for this student population.

## 2. Literature Review

### 2.1 Metacognitive Regulation and the Demands of Critical Writing

Metacognition is conventionally divided into knowledge about cognition and regulation of cognition (Flavell, 1979; Nelson & Narens, 1990). This study focuses on regulation because it is the component instruction and design can influence. It is modelled as a cycle in which learners plan, monitor progress against goals, and evaluate outcomes to inform later attempts (Pintrich, 2000; Winne & Hadwin, 1998; Zimmerman, 2000, 2002). The design implication is direct: support for one phase cannot substitute for support at another.

Writing makes this cycle demanding. Process models treat writing as recursive rather than linear, with planning, translating, and reviewing under executive control (Flower & Hayes, 1981), and distinguish knowledge-telling from knowledge-transforming, in which writing reshapes understanding (Bereiter & Scardamalia, 1987). Academic writing depends on the second mode: writers must hold a rhetorical goal in view, detect gaps between the current text and that goal, and act on them. Critical writing intensifies the demand because coordinating claim, evidence, warrant, and rebuttal (Toulmin, 1958) develops unevenly and often late (Kuhn, 1991), while expert consensus places self-regulation among the core skills of critical thinking (Facione, 1990). Monitoring and correcting one's reasoning is therefore constitutive of the competence, not ancillary to it.

Empirical work supports this position: strategy instruction targeting planning, monitoring, and evaluation improves writing performance and motivation (Han, 2024; L. S. Teng & Zhang, 2020); metacognitive strategy use transfers across first- and second-language writing (Xu & Zhu, 2024); self-regulated learning interventions in online and blended settings produce reliable gains (Guntur & Purnomo, 2024; Luo & Zhou, 2024); and critical thinking has responded to instructional design with Chinese and other second-language learners (Liu & Wang, 2024; X. Song et al., 2025). Two consequences follow. A needs assessment must ask about planning, monitoring, and evaluation separately rather than treating metacognitive support as one commodity; and because the cycle is integrated, items addressing different phases should cohere empirically, an expectation RQ1 tests.

### 2.2 Digital Learning Environments and the Asymmetry of Regulatory Support

The idea that an environment can carry part of a learner's regulatory burden descends from scaffolding: a more capable other, or designed artefact, temporarily assumes functions the learner cannot yet perform alone and withdraws assistance as competence develops (Vygotsky, 1978; Wood et al., 1976). Digital environments extend this logic by embedding prompts, planning templates, progress indicators, and feedback into students' tools. Adaptive support has become a design agenda (Khalil et al., 2024), agent scaffolding works best when aligned with learners' subgoals (Dever et al., 2024), and reviews show that technology can support regulation directly and indirectly (Michael et al., 2025), although pre-structured environments risk performing regulation for learners rather than developing it (Herold & Seufert, 2025).

Recent international work extends this issue: embedded prompts, dashboards, feedback, and adaptive scaffolds can support regulation, yet their effects depend on timing, interpretation, and learner use. This study therefore treats Chinese evidence as part of a broader debate about how digital environments distribute regulation between learner and system.

Support is uneven across the regulatory cycle. Monitoring is well served: progress bars, completion indicators, and dashboards make current state visible. Planning may be less provisioned because helping learners construct goals before work begins is harder to instrument than tracking completed work; evaluation may be harder still because it requires support for judging quality against criteria. The prevailing paradigm predicts outcomes rather than describing process, yet regulation depends on actionable process information.

Feedback is a main route through which digital environments influence writing. Meta-analytic and review evidence reports gains from automated and AI-generated feedback on second-language writing performance and related outcomes (Fleckenstein et al., 2023; Shi & Aryadoust, 2024), while uptake depends on metacognitive awareness (M. F. Teng, 2025). Automated feedback is often evaluative and prescriptive rather than descriptive of the writer's process; in Hattie and Timperley's (2007) terms, it operates more often at the task level than at the process and self-regulation levels that support independent revision; telling a writer what to change differs from helping a writer see how they worked. This asymmetry generates a directional expectation for RQ2: if planning and evaluation are under-provisioned relative to monitoring, students should register the shortfall.

### 2.3 Learner Beliefs, Expressed Demand, and the Assessment of Needs

Self-efficacy—belief in one’s capacity to execute actions required for an outcome—shapes effort, persistence, and goal setting (Bandura, 1977, 1997), and in writing has long been linked to performance and strategy use (Pajares, 2003; Zhang & Zhang, 2024). Its measurement is contested: recent scales and reviews argue that the construct is multidimensional and may not generalise across tasks (Collado et al., 2023; J. Wang et al., 2026). For instrument design, a few writing-competence items cannot be assumed to form a unitary construct.

A second strand concerns how expressed need should be read. Academic help-seeking is theorised not as dependence but as a self-regulatory strategy: learners must monitor comprehension, recognise insufficiency, identify an adequate source, and act (Karabenick & Newman, 2006; Newman, 1994; Won & Chang, 2024). Demand for support may therefore indicate regulatory awareness rather than its absence, meaning better provisioned or more capable learners might report *more* need, not less.

These questions are culturally situated. Chinese higher education often combines strong instructional structure with developing learner autonomy, so students may value guidance while still needing support for planning, monitoring, and evaluation. This context makes it important to ask what learners themselves perceive as missing.

Needs assessment grounds design in context. Expert review can establish item relevance, but exploratory factor analysis is still needed to test whether students distinguish provision, support need, and competence beliefs.

## 3. Methodology

### 3.1 Research Design

This cross-sectional quantitative survey established the empirical basis for a digital learning environment model (DLEM) to enhance metacognitive regulation in critical writing. It formed the context-analysis component of Phase 1 within a Type 2 design and development research framework (Richey & Klein, 2007). The design suited the descriptive aims: to characterise current provision, quantify metacognitive-support needs, and test subgroup differences. No intervention or temporal comparison was involved.

### 3.2 Setting and Participants

The study was conducted at two private undergraduate institutions in Hechuan District, Chongqing, China: Paisi College of Chongqing Technology and Business University, and Chongqing College of Humanities, Science and Technology. Both shared comparable teaching-management principles, student profiles, and critical-writing provision, allowing aggregation while retaining institution as an analytic contrast. Both are private undergraduate colleges, a sector that may differ from public universities in student intake, resourcing, and digital infrastructure; this bounds the population to which the findings can be generalised (Section 6).

Participants were 400 second-year undergraduates, 200 from each institution. They were targeted because critical writing is taught at this stage and they had enough experience with the institutional digital environment to evaluate it. Of the respondents, 188 (47.0%) were male, 212 (53.0%) female, 318 (79.5%) reported prior online-learning experience, and 82 (20.5%) did not.

Sample size was set a priori. Factor-analytic conventions recommend a 10:1 subject-to-item ratio and at least 300 cases (Comrey & Lee, 1992; MacCallum et al., 1999); with 15 items,  $N = 400$  gives 26.7:1. G\*Power 3.1 (Faul et al., 2007) indicated .98 power to detect  $d = 0.40$  with two groups of 200 at  $\alpha = .05$ .

### 3.3 Instrument

The Digital Learning Environment Teaching Needs Survey was developed for this study. Part 1 recorded institution, year, gender, and prior online-learning experience. Part 2 contained 15 five-point Likert items (1 = *strongly disagree* to 5 = *strongly agree*) across three conceptual blocks: current digital learning environment use (Q6–Q10), metacognitive regulation support needs (Q11–Q15), and critical writing status and needs (Q16–Q20). Item wording appears in Table 1.

Item development followed standard design-and-development procedures: the framework was derived from the literature, sub-dimensions were specified, an item pool was drafted and reviewed by the supervisor, and the revised instrument was submitted to an expert panel.

Five experts independently rated item relevance. Eleven items had I-CVI = 1.00 and four had I-CVI = 0.80 (Table S1), all meeting the 0.78 criterion for five-expert panels (Lynn, 1986). The S-CVI/Ave was 0.95, exceeding the 0.90 threshold for scale-level content validity (Polit & Beck, 2006). Wording revisions followed expert feedback.

### 3.4 Ethical Considerations

The protocol was approved by the Center for Ethics in Human Research, Khon Kaen University (HE693082; IRB00012791; FWA00003418), and both institutions granted permission. Participants received written information,

gave consent, could decline items or withdraw without academic consequence, and provided no identifying information. Questionnaires used anonymous codes (S001–S400) and sealed collection boxes.

### 3.5 Data Collection and Preparation

The paper questionnaire was administered in intact classes over three weeks by non-teaching research assistants and took about 20 minutes. A random subset was independently re-entered to check accuracy. Screening examined out-of-range values, skewness, and kurtosis against maximum-likelihood criteria ( $|\text{skew}| < 1.0$ ;  $|\text{kurtosis}| < 2.0$ ).

Response-set bias was assessed using Harman's single-factor test (Podsakoff et al., 2003) and by comparing the inter-item correlation structure before and after within-person centring.

Missing data were minimal: 24 of 6,000 responses (0.4%), with no item exceeding 1.0%, and were consistent with missing completely at random. Listwise deletion was used for factor analysis ( $n = 378$ ) and available-case analysis elsewhere; no imputation was performed.

### 3.6 Data Analysis

Statistical significance was evaluated at  $\alpha = .05$  unless otherwise stated.

#### 3.6.1 Dimensionality and Reliability (RQ1)

Because the instrument was new, exploratory factor analysis was used. Factorability was assessed with KMO and Bartlett's test, and factor number by parallel analysis (Horn, 1965). Factors were extracted using maximum likelihood and promax rotation. Items were retained if they loaded  $\geq .40$  on a primary factor, did not cross-load  $\geq .30$ , and achieved communality  $\geq .30$ ; failing items were removed and the solution re-estimated. Internal consistency was evaluated using Cronbach's  $\alpha$ , McDonald's  $\omega$ , and, for two-item scales, the Spearman–Brown coefficient with inter-item correlation (Eisinga et al., 2013).

#### 3.6.2 Need–provision discrepancy (RQ2)

A paired-samples  $t$ -test compared Support Need and DLE Provision, reporting Cohen's  $d_z$  and the proportion of students with a positive discrepancy. Item priorities used within-person centring, subtracting each respondent's grand mean from each item score to reduce scale-use bias and show *relative* priority.

#### 3.6.3 Subgroup variation and profile analysis (RQ3)

Institution, gender, and prior online-learning differences were tested using independent-samples  $t$ -tests on each retained scale. Latent profile analysis used standardised scale scores, Gaussian mixture models with diagonal covariance, and 50 random initialisations. One- to five-profile models were compared using Akaike and Bayesian information criteria, entropy, and smallest-class size; entropy  $\geq .80$  marked confident assignment (Celeux & Soromenho, 1996). Multinomial logistic regression tested whether background variables predicted membership.

#### 3.6.4 Equivalence testing

Because subgroup analyses yielded null results, two one-sided tests (TOST) distinguished absence of evidence from evidence of absence (Lakens, 2017). The smallest effect size of interest was set a priori at  $d = \pm 0.40$ : about one-third of a scale point on a five-point scale with standard deviations near 0.85, below which subgroup-specific pathways would not be warranted. Equivalence was concluded when the 90% confidence interval fell entirely within the bounds. A sensitivity analysis used  $d = \pm 0.30$ , with achieved power reported for each comparison.

### 3.7 Reproducibility

A scripted pipeline regenerated all reported tables and figures from the raw item-level data. Thresholds were predeclared in code, item removal was automated, fixed seeds controlled stochastic steps, and repeated execution reproduced all numerical output.

Analyses used Python 3.12.3 with numpy, pandas, scipy, scikit-learn, statsmodels, factor\_analyzer, and matplotlib. The de-identified dataset, analysis script, codebook, console log, and environment specification are available from the corresponding author on reasonable request.

## 4. Results

### 4.1 Dimensional Structure of Students' Perceptions (RQ1)

#### 4.1.1 Screening Outcomes

Table 1 reports item-level statistics. Means ranged from 2.91 to 3.64; skewness and kurtosis were within maximum-likelihood criteria. Respondents discriminated across items rather than endorsing a single response set.

Table 1. Item-level descriptive statistics for the 15 survey items, by designed block

| Item | Description                                   | <i>n</i> | <i>M</i> | <i>SD</i> | Skew  | Kurtosis |
|------|-----------------------------------------------|----------|----------|-----------|-------|----------|
| Q6   | Frequent use of digital learning tools        | 399      | 3.29     | 1.16      | -0.21 | -0.81    |
| Q7   | Platform functions are comprehensive          | 398      | 3.05     | 1.24      | -0.01 | -1.03    |
| Q8   | School provides sufficient digital resources  | 399      | 3.19     | 1.16      | -0.15 | -0.79    |
| Q9   | Technical support is timely and effective     | 397      | 3.01     | 1.11      | -0.01 | -0.71    |
| Q10  | Overall satisfaction with the DLE             | 399      | 3.21     | 1.17      | -0.12 | -0.81    |
| Q11  | Need tools for planning and goal setting      | 398      | 3.57     | 1.15      | -0.34 | -0.93    |
| Q12  | Need tools to monitor writing progress        | 399      | 3.58     | 1.12      | -0.46 | -0.51    |
| Q13  | Need tools for self-evaluation and reflection | 398      | 3.53     | 1.07      | -0.32 | -0.67    |
| Q14  | Need structured guidance and model references | 400      | 3.64     | 1.13      | -0.47 | -0.55    |
| Q15  | Peer feedback and collaboration are important | 398      | 3.32     | 1.12      | -0.25 | -0.72    |
| Q16  | Encounter difficulty in critical writing      | 400      | 3.20     | 1.20      | -0.02 | -0.98    |
| Q17  | Know how to construct a persuasive argument   | 398      | 2.91     | 1.27      | 0.14  | -1.00    |
| Q18  | Can evaluate and use evidence effectively     | 399      | 2.93     | 1.26      | 0.11  | -0.99    |
| Q19  | Want more guidance and practice opportunities | 396      | 3.54     | 1.21      | -0.39 | -0.84    |
| Q20  | DLE can improve critical writing ability      | 398      | 3.53     | 1.18      | -0.35 | -0.82    |

*Note.* *N* = 400; *n* varies by item owing to missing responses. Response scale 1–5. Designed blocks: Q6–Q10 = current digital learning environment (DLE) use; Q11–Q15 = metacognitive regulation support needs; Q16–Q20 = critical writing status and needs. Skewness within  $\pm 1.0$  and kurtosis within  $\pm 2.0$  indicate distributions suitable for maximum-likelihood estimation.

#### 4.1.2 Factorability and Factor Retention

Sampling adequacy was adequate overall, KMO = .861, and Bartlett's test of sphericity was significant,  $\chi^2(105) = 1710.7$ ,  $p < .001$ . Item-level KMO values ranged from .58 to .92, the two lowest belonging to Q17 and Q18, which shared less variance with the remaining pool.

Parallel analysis supported three factors: the first three observed eigenvalues (4.58, 2.12, 1.60) exceeded their random criteria (1.42, 1.33, 1.26), whereas the fourth (0.94) fell below its criterion (1.20). One item was then removed by the retention procedure: Q16 ("I have encountered difficulties in critical writing") achieved a primary loading of .25 and a communality of .12. Re-estimation on the reduced set satisfied all criteria at the first iteration and did not disturb any other item's assignment.

#### 4.1.3 The three-factor Solution

The final solution (Table 2) accounted for 46.9% of total variance. Simple structure was achieved: every item loaded  $\geq .51$  on its primary factor, the largest secondary loading anywhere in the matrix was .27 (Q20 on the provision factor), and communalities ranged from .34 (Q20) to .62 (Q17).

Table 2. Rotated pattern matrix of the three-factor solution, showing which items students grouped together (maximum likelihood, promax; *n* = 378)

| Item | F1 Support Need | F2 DLE Provision | F3 CW Self-Efficacy | <i>h</i> <sup>2</sup> |
|------|-----------------|------------------|---------------------|-----------------------|
| Q11  | <b>.77</b>      | -.08             | .06                 | .61                   |
| Q12  | <b>.75</b>      | -.09             | .07                 | .57                   |
| Q14  | <b>.73</b>      | -.05             | .05                 | .54                   |
| Q15  | <b>.65</b>      | -.02             | .08                 | .43                   |
| Q13  | <b>.63</b>      | .07              | -.04                | .40                   |
| Q19  | <b>.58</b>      | .04              | -.19                | .37                   |
| Q20  | <b>.51</b>      | .27              | -.09                | .34                   |
| Q6   | -.04            | <b>.70</b>       | -.02                | .49                   |
| Q7   | .00             | <b>.70</b>       | .05                 | .49                   |
| Q10  | -.08            | <b>.69</b>       | -.03                | .49                   |
| Q9   | -.01            | <b>.65</b>       | .04                 | .42                   |
| Q8   | .06             | <b>.61</b>       | .04                 | .37                   |
| Q17  | .06             | .02              | <b>.78</b>          | .62                   |
| Q18  | .00             | .04              | <b>.65</b>          | .43                   |

*Note.* Primary loadings in bold. Q16 removed prior to the final extraction. *h*<sup>2</sup> = communality. SS loadings = 3.12 / 2.34 / 1.10; cumulative variance = 46.9%. Retention required a primary loading  $\geq .40$ , no cross-loading  $\geq .30$ , and communality  $\geq .30$ . F1 items express demand for support from the environment; F2 items appraise current provision; F3 items appraise one's own argument competence. CW = critical writing.

The recovered structure departed from the a priori design. The first block (Q6–Q10) emerged intact as Factor 2 Perceived DLE Provision. The second (Q11–Q15) held together as Factor 1 Perceived Support Need and absorbed Q19 and Q20, yielding a seven-item scale. Q17 and Q18 formed Factor 3 Critical Writing Self-Efficacy.

Q19 and Q20 migrated coherently: both express demands directed at the learning environment and align more closely with support need than competence. The third designed block had conflated writing-ability appraisal with desired environmental support. Q16 aligned with no factor strongly enough to retain, correlating  $+0.33$  with support need,  $+0.22$  with provision, and  $-0.19$  with self-efficacy—modest, shared variance with all three, membership of none.

Factor intercorrelations followed this pattern. Support Need correlated positively with DLE Provision ( $\phi = .42$ ) and negatively with Self-Efficacy ( $\phi = -.20$ ), while Provision and Self-Efficacy were unrelated ( $\phi = -.10$ ). Observed composite correlations showed the same pattern:  $r = .33$  ( $p < .001$ ),  $r = -.10$  ( $p = .048$ ), and  $r = .01$  ( $p = .847$ ).

#### 4.1.4 Reliability

Internal consistency was acceptable for Support Need and DLE Provision. The two-item Critical Writing Self-Efficacy scale showed modest reliability, so it is treated as a focused indicator rather than a full self-efficacy measure.

Table 3. Composition, reliability, and descriptive statistics of the three retained scales

| Scale                             | Items             | <i>k</i> | $\alpha$ | $\omega$ | <i>M</i> | <i>SD</i> |
|-----------------------------------|-------------------|----------|----------|----------|----------|-----------|
| F1 Perceived Support Need         | Q11–Q15, Q19, Q20 | 7        | .852     | .846     | 3.53     | 0.83      |
| F2 Perceived DLE Provision        | Q6–Q10            | 5        | .801     | .802     | 3.15     | 0.87      |
| F3 Critical Writing Self-Efficacy | Q17, Q18          | 2        | .667     | .681     | 2.92     | 1.10      |

*Note.* For the two-item self-efficacy scale the Spearman–Brown coefficient is reported in place of  $\alpha$ ; inter-item  $r = .50$ .  $\omega$  = McDonald's omega. Scale scores are item means on the 1–5 response scale, on which 3.0 is the midpoint; only Critical Writing Self-Efficacy falls below it.

#### 4.2 The Need–provision Discrepancy and Its Composition (RQ2)

##### 4.2.1 Magnitude and Prevalence of the Discrepancy

Perceived Support Need ( $M = 3.53$ ,  $SD = 0.83$ ) exceeded Perceived DLE Provision ( $M = 3.15$ ,  $SD = 0.87$ ) by 0.38 scale points, 95% CI [0.28, 0.48]. The difference was significant,  $t(399) = 7.73$ ,  $p < .001$ ,  $d_z = 0.39$ .

The distribution clarifies the pattern (Figure 1). Of 400 students, 262 (65.5%) reported need above provision, 126 (31.5%) the reverse, and 12 (3.0%) no gap. Among those with unmet need, 98 (24.5% of the full sample) showed a gap exceeding one scale point and 83 (20.8%) a gap between half and one point. Scores were close to symmetric ( $Mdn = 0.40$ , skew =  $-0.18$ ), indicating a distributional shift rather than an effect driven by extreme respondents.

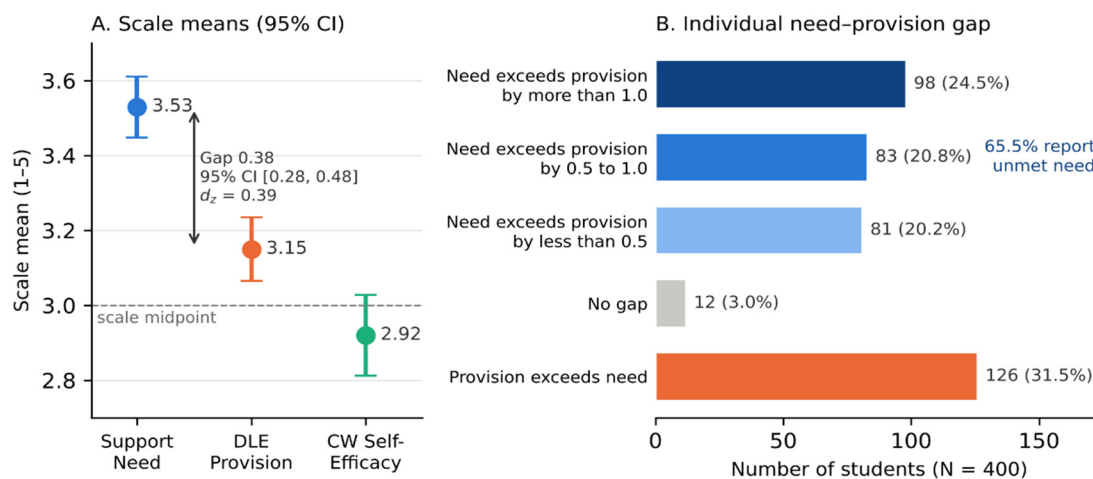


Figure 1. Need–provision discrepancy at the scale level and for individual students (N = 400)

*Note.* Panel A shows scale means with 95% confidence intervals; the dashed line marks the scale midpoint, and the arrow marks the mean need–provision gap from the paired  $t$ -test. Panel B classifies each student by the difference between their Support Need and DLE Provision scores; the three blue bars together make up the 65.5% of students reporting unmet need.

Absolute levels add context. Agreement (responses of 4 or 5) with provision items ranged from 34.0% to 45.6%, with 36.5% scoring below the provision-scale midpoint; need items drew agreement between 47.2% and 56.5%. The

discrepancy therefore appears to reflect generic provision alongside unmet writing-specific requirements, not absent digital infrastructure.

#### 4.2.2 Item-level Priorities

Relative priorities after within-person centring are shown in Figure 2. The ordering divides sharply into two blocks with almost no overlap.

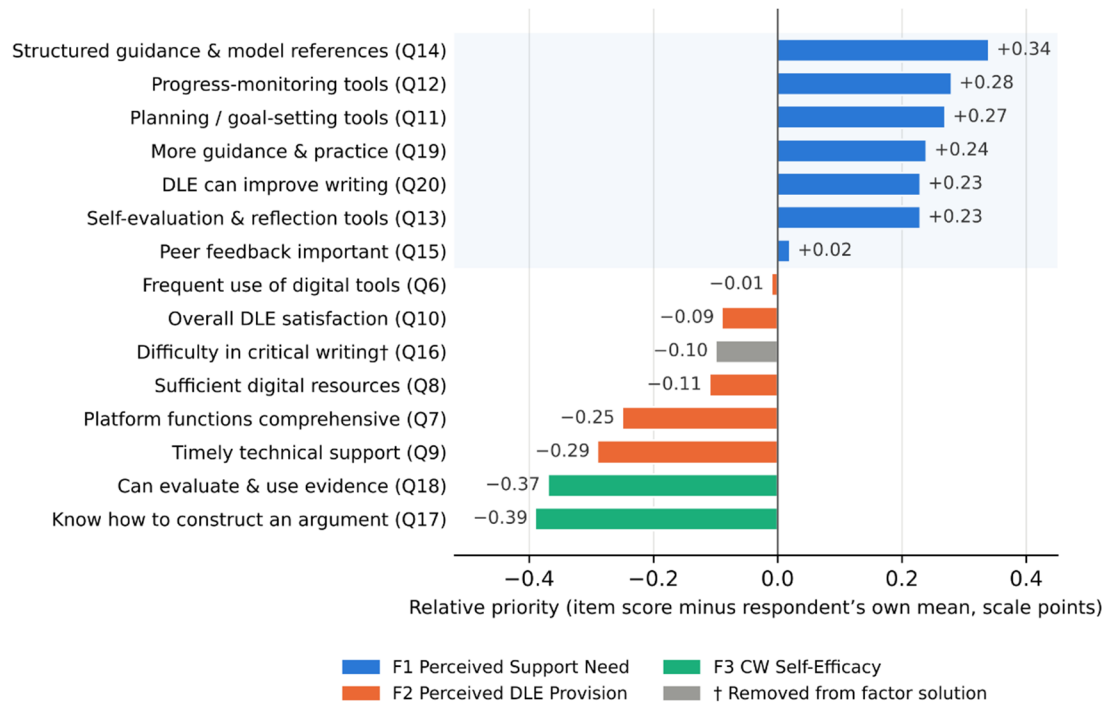


Figure 2. Relative priority of the 15 survey items after within-person centring, coloured by factor membership (N = 400)

*Note.* Positive values indicate items rated above a respondent's own average across all items; negative values indicate below-average priority. This shows relative priority rather than absolute agreement.

The six highest priorities, from +0.34 to +0.23, were structured guidance and model references, progress-monitoring tools, planning and goal-setting tools, more guidance and practice, belief in the potential of digital environments, and self-evaluation and reflection tools. All request scaffolding of the metacognitive cycle. The lowest priorities were the two self-efficacy statements—knowing how to construct an argument (−0.39) and being able to evaluate evidence (−0.37)—followed by timely technical support (−0.29) and comprehensive platform functions (−0.25).

Peer feedback (+0.02) was neutral, the only need item outside the priority group. This is notable given the emphasis on collaborative learning in this literature. The result may reflect examination-oriented settings in which learner autonomy and peer-based learning remain under-developed (Su et al., 2023), or limited exposure to structured peer review; this cannot be settled here and matters because collaborative features are easy to add but may not be students' main priority.

#### 4.2.3 The position of Self-efficacy

Critical Writing Self-Efficacy was the lowest-scoring construct ( $M = 2.92$ ) and the only scale below the midpoint, with 46.0% of students below 3.0. About one third agreed that they knew how to construct a persuasive argument (32.7%) or could evaluate and use evidence (33.1%); about 40% disagreed with each, the highest disagreement rates in the instrument.

Self-efficacy correlated negatively with support need ( $r = -.10, p = .048$ ): students who felt least capable expressed the greatest demand for scaffolding. The association is weak and is not interpreted further, though stronger at the factor level ( $\phi = -.20$ ); the removed item Q16 behaved similarly, correlating  $-.18$  with argument-construction competence and  $+.41$  with desire for further guidance. Self-efficacy was unrelated to perceived provision ( $r = .01, p = .847$ ): students who rated the digital environment highly felt no more capable as critical writers than those who rated it poorly.

### 4.3 Uniformity of Needs Across Subgroups (RQ3)

#### 4.3.1 Direct Comparisons

No scale differed significantly by any background variable (Table 4). The largest effect was for Support Need between

institutions ( $M_{\text{diff}} = 0.15$ ,  $d = 0.18$ ,  $p = .077$ ); the other eight comparisons had  $|d| \leq 0.18$  and  $p$  values from .157 to .905. Provision and self-efficacy were near-identical across institutions (3.17 versus 3.13, and 2.90 versus 2.94), and the mean need–provision gap was similar in both settings, 0.43 and 0.33.

Table 4. Subgroup comparisons with equivalence tests (TOST) showing that needs did not differ meaningfully by institution, gender, or prior online experience; primary bound  $d = \pm 0.40$

| Comparison                    | Scale            | $n_1 / n_2$ | $M_{\text{diff}}$ | $d$   | $p$ (NHST) | 90% CI         | $p$ (TOST) | Equivalent |
|-------------------------------|------------------|-------------|-------------------|-------|------------|----------------|------------|------------|
| Institution (1 vs 2)          | Support Need     | 200 / 200   | +0.15             | +0.18 | .077       | [+0.01, +0.28] | .013       | Yes        |
|                               | DLE Provision    | 200 / 200   | +0.05             | +0.05 | .596       | [−0.10, +0.19] | < .001     | Yes        |
|                               | CW Self-Efficacy | 200 / 200   | −0.05             | −0.04 | .666       | [−0.23, +0.13] | < .001     | Yes        |
| Gender (male vs female)       | Support Need     | 188 / 212   | +0.04             | +0.05 | .634       | [−0.10, +0.18] | < .001     | Yes        |
|                               | DLE Provision    | 188 / 212   | −0.06             | −0.07 | .517       | [−0.20, +0.09] | .001       | Yes        |
|                               | CW Self-Efficacy | 188 / 212   | +0.01             | +0.01 | .905       | [−0.17, +0.20] | < .001     | Yes        |
| Online experience (no vs yes) | Support Need     | 82 / 318    | −0.13             | −0.15 | .220       | [−0.29, +0.04] | .023       | Yes        |
|                               | DLE Provision    | 82 / 318    | +0.13             | +0.15 | .230       | [−0.05, +0.31] | .022       | Yes        |
|                               | CW Self-Efficacy | 82 / 318    | −0.19             | −0.18 | .157       | [−0.42, +0.03] | .035       | Yes        |

*Note.* Equivalence is concluded when the 90% CI falls entirely within the bounds. Sensitivity results at  $d = \pm 0.30$  are given in the text. Positive differences indicate a higher mean in the first-listed group. NHST = two-tailed independent-samples t-test; TOST  $p$  = the larger of the two one-sided  $p$  values.

#### 4.3.2 Exploratory Profile Analysis

Latent profile analysis found no well-separated need profiles. Although fit indices varied, entropy was low and profile differences were descriptive rather than actionable.

#### 4.3.3 Equivalence Testing

All nine equivalence tests were significant at the primary bound of  $d = \pm 0.40$ , with  $p$  values from < .001 to .035 (Table 4); in each case the 90% confidence interval of the mean difference fell entirely within the bounds. Achieved power to detect  $d = 0.40$  was .98 for the institution and gender comparisons and .90 for online experience.

The sensitivity analysis at  $d = \pm 0.30$  produced a more qualified picture. All three gender comparisons remained equivalent ( $p = .002$  to .010), as did two institution comparisons (provision  $p = .007$ , self-efficacy  $p = .005$ ), while institutional differences in support need did not ( $p = .110$ ). None of the online-experience comparisons reached equivalence at the stricter bound ( $p = .112$  to .158); this follows from the unbalanced group sizes (82 versus 318), which reduced power to .68, since observed effect sizes were  $|d| \leq 0.18$  throughout.

### 5. Discussion

#### 5.1 RQ1: The Dimensional Structure of Students' Perceptions

Students distinguished three constructs the instrument was not designed to separate: environmental provision, desired support, and perceived writing capability. Read against Bandura's (1977, 1997) account, the separation is coherent: efficacy beliefs concern what one can execute, whereas desired environmental support is an outward-directed outcome expectation. Students appear to hold these apart even when the questionnaire does not.

The solution left one issue unresolved: planning, monitoring, and evaluation items loaded on a single support-need factor. Students did not separate regulatory phases as theory does, perhaps because they experience these supports as one practical need.

Designers should therefore avoid assuming that separate platform features will be experienced as separate supports. Future instruments should include more items per regulatory phase.

#### 5.2 RQ2: The Discrepancy Between Need and Provision

The documented gap indicates adequate but unremarkable infrastructure alongside unmet writing-regulation needs. The priority is not more technology in general, but support for planning, monitoring, and evaluation.

One element of that ordering was unexpected. If digital environments over-serve monitoring relative to planning and evaluation, monitoring should rank lower; instead, it ranked second. Students may distinguish activity visibility, which dashboards supply, from interpretive support for acting on displayed information, consistent with reviews finding dashboards weakly grounded in regulation models (Cabral et al., 2025; Matcha et al., 2020; Paulsen & Lindsay, 2024).

Alternatively, the single monitoring item may have been read as a general request for feedback.

Peer feedback was neutral despite its prominence in the literature. It may need staged introduction after students have practised self-evaluation and evidence-based revision.

The positive correlation between provision and need may reflect regulatory awareness: students who recognise useful support also see what additional support could help them.

### *5.3 RQ3: Uniformity of Needs Across Subgroups*

Expressed needs did not differ by institution, gender, or prior online-learning experience, and equivalence tests support a stronger conclusion than failure to detect difference: every comparison fell within  $d = \pm 0.40$ . A model for this population may not require pathways tailored to these characteristics, making a universal scaffolding architecture the more defensible starting point. This conclusion is unavailable from non-significant *t*-tests alone, which remain compatible with real differences the study was underpowered to detect (Ponce-Renova, 2022). The profile analysis was consistent: no solution achieved acceptable class separation, and no background variable predicted membership.

This differs from profile studies of regulatory capability because this study profiled expressed demand rather than learner competence.

Uniform demand does not imply uniform benefit. Future work should test whether a common scaffold helps all learners equally once actual performance and log data are observed.

### *5.4 Implications for the Design of Digital Learning Environments*

The findings suggest a universal core: structured guidance, model references, planning prompts, progress interpretation, and self-evaluation tools attached to writing tasks rather than generic platform use.

At the same time, the system should allow optional intensification. Learners with low confidence or weak engagement could receive extra examples, planning prompts, or reflection questions without creating fixed pathways by gender or institution.

Peer feedback and adaptive release should be staged. Students first need individual scaffolds for planning, monitoring, and evaluation; collaborative and adaptive functions can then build on that foundation.

## **6. Limitations**

Several constraints bound what can be claimed from these data.

**Design and inference.** The study is cross-sectional. It documents perceptions at one point in time and supports no statement about what produces them or about what would follow from changing them; the associations reported are correlational.

**Measurement.** The study used self-reports, so it captures perceived need rather than demonstrated need. The two-item self-efficacy scale is narrow and modest in reliability; future work should add items, writing samples, interviews, or log data.

**Analysis.** The factor structure was exploratory and has not been confirmed on an independent sample; it is a hypothesis about the instrument, not established measurement. The profile labels are not validated learner types. The analysis was not pre-registered, and some Type I error inflation across tests cannot be ruled out.

**Generalisability.** Participants came from two private institutions in one district of Chongqing and one course context. Findings should be replicated in public universities, other regions, other years, and different writing courses.

## **7. Conclusion**

This study found that Chinese undergraduates perceived a clear metacognitive support gap in digital critical-writing environments. Support need exceeded provision, with strongest priorities for structured guidance, progress monitoring, planning, and self-evaluation. Needs were equivalent across the measured subgroups, supporting a common scaffolding core with optional individual intensification. The findings offer a context-grounded basis for designing DLEs that support regulation rather than only deliver content.

## **Acknowledgments**

The authors thank the five experts who reviewed the survey instrument, the students who took part, and the two participating institutions for their cooperation.

## **Author contributions**

Jinjin Jiang was responsible for the study design, data collection, data analysis, and drafting of the manuscript. Parama Kwangmuang supervised the study, contributed to the research design and interpretation of the findings, and revised the manuscript. Both authors read and approved the final manuscript.

**Funding**

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

**Competing interests**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Informed consent**

Obtained.

**Ethics approval**

The Publication Ethics Committee of the Redfame Publishing.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

**Provenance and peer review**

Not commissioned; externally double-blind peer reviewed.

**Data availability statement**

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

**Data sharing statement**

No additional data are available.

**Open access**

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

**Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

**References**

- Azevedo, R., & Hadwin, A. F. (2005). Scaffolding self-regulated learning and metacognition – Implications for the design of computer-based scaffolds. *Instructional Science*, 33(5–6), 367–379. <https://doi.org/10.1007/s11251-005-1272-9>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
- Bannert, M., & Reimann, P. (2012). Supporting self-regulated hypermedia learning through prompts. *Instructional Science*, 40(1), 193–211. <https://doi.org/10.1007/s11251-011-9167-4>
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Lawrence Erlbaum Associates.
- Bråten, I., Haverkamp, Y. E., Latini, N., & Strømsø, H. I. (2023). Measuring multiple-source based academic writing self-efficacy. *Frontiers in Psychology*, 14, 1212567. <https://doi.org/10.3389/fpsyg.2023.1212567>
- Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, 27, 1–13. <https://doi.org/10.1016/j.iheduc.2015.04.007>
- Cabral, L., Pinto, R., & Gonçalves, G. (2025). AI-powered learning analytics dashboards: A systematic review of applications, techniques, and research gaps. *Discover Education*, 4(1), Article 525. <https://doi.org/10.1007/s44217-025-00964-y>
- CAST. (2024). *Universal design for learning guidelines version 3.0*. <https://udlguidelines.cast.org>
- Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*, 13(2), 195–212. <https://doi.org/10.1007/BF01246098>
- Collado, S., Fidalgo, C., Rodríguez-Rey, R., & Sorrel, M. A. (2023). Development and validation of the self-efficacy for writing and defending academic texts scale. *Psicología Educativa*, 29(1), 91–99. <https://doi.org/10.5093/psed2022a15>
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Lawrence Erlbaum Associates.

- Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268. [https://doi.org/10.1207/S15327965PLI1104\\_01](https://doi.org/10.1207/S15327965PLI1104_01)
- Dever, D. A., Wiedbusch, M. D., Romero, S. M., & Azevedo, R. (2024). Investigating pedagogical agents' scaffolding of self-regulated learning in relation to learners' subgoals. *British Journal of Educational Technology*, 55(4), 1290–1308. <https://doi.org/10.1111/bjet.13432>
- Devolder, A., van Braak, J., & Tondeur, J. (2012). Supporting self-regulated learning in computer-based learning environments: Systematic review of effects of scaffolding in the domain of science education. *Journal of Computer Assisted Learning*, 28(6), 557–573. <https://doi.org/10.1111/j.1365-2729.2011.00476.x>
- Eisinga, R., te Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health*, 58(4), 637–642. <https://doi.org/10.1007/s00038-012-0416-3>
- Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction* (The Delphi Report). American Philosophical Association.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, 34(10), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Fleckenstein, J., Liebenow, L. W., & Meyer, J. (2023). Automated feedback and writing: A multi-level meta-analysis of effects on students' performance. *Frontiers in Artificial Intelligence*, 6, 1162454. <https://doi.org/10.3389/frai.2023.1162454>
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387. <https://doi.org/10.2307/356600>
- Guntur, M., & Purnomo, Y. W. (2024). A meta-analysis of self-regulated learning interventions studies on learning outcomes in online and blended environments. *Online Learning*, 28(3), 563–584. <https://doi.org/10.24059/olj.v28i3.4025>
- Guo, L. (2022). Using metacognitive prompts to enhance self-regulated learning and learning outcomes: A meta-analysis of experimental studies in computer-based learning environments. *Journal of Computer Assisted Learning*, 38(3), 811–832. <https://doi.org/10.1111/jcal.12650>
- Han, L. (2024). Metacognitive writing strategy instruction in the EFL context: Focus on writing performance and motivation. *SAGE Open*, 14(2). <https://doi.org/10.1177/21582440241257081>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Herold, A., & Seufert, T. (2025). Who regulates? Effects of scaffolding in system- and self-regulated learning. *Frontiers in Psychology*, 16, 1543024. <https://doi.org/10.3389/fpsyg.2025.1543024>
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations* (2<sup>nd</sup> ed.). Sage.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179–185. <https://doi.org/10.1007/BF02289447>
- Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, 19(4), 509–539. <https://doi.org/10.1007/s10648-007-9054-3>
- Karabenick, S. A., & Newman, R. S. (Eds.). (2006). *Help seeking in academic settings: Goals, groups, and contexts*. Lawrence Erlbaum Associates.
- Kember, D. (2000). Misconceptions about the learning approaches, motivation and study practices of Asian students. *Higher Education*, 40(1), 99–121. <https://doi.org/10.1023/A:1004036826490>
- Khalil, M., Wong, J., Wasson, B., & Paas, F. (2024). Adaptive support for self-regulated learning in digital learning environments. *British Journal of Educational Technology*, 55(4), 1281–1289. <https://doi.org/10.1111/bjet.13479>
- Kuhn, D. (1991). *The skills of argument*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511571350>
- Lakens, D. (2017). Equivalence tests: A practical primer for *t* tests, correlations, and meta-analyses. *Social Psychological*

- and *Personality Science*, 8(4), 355–362. <https://doi.org/10.1177/1948550617697177>
- Li, C., Lee, I., & Mao, Z. (2026). Synthesizing empirical research on learner metacognition in second language writing. *International Journal of Applied Linguistics*. <https://doi.org/10.1111/ijal.70123>
- Li, L., Chen, S., Tang, P., Lv, K., Li, P., & Xu, Y. (2025). Unlocking deep learning: How self-regulated learning shapes Chinese students' online course experiences. *SAGE Open*, 15(3). <https://doi.org/10.1177/21582440251365390>
- Li, Z., Wu, J., Thurston, A., & O'Keeffe, J. (2026). A systematic review of second language (L2) student writers' metacognitive experiences. *British Educational Research Journal*, 52(3), 2281–2315. <https://doi.org/10.1002/berj.70109>
- Littlewood, W. (1999). Defining and developing autonomy in East Asian contexts. *Applied Linguistics*, 20(1), 71–94. <https://doi.org/10.1093/applin/20.1.71>
- Liu, W., & Wang, Y. (2024). The effects of using AI tools on critical thinking in English literature classes among EFL learners: An intervention study. *European Journal of Education*, 59(4), Article e12804. <https://doi.org/10.1111/ejed.12804>
- Luo, R.-Z., & Zhou, Y.-L. (2024). The effectiveness of self-regulated learning strategies in higher education blended learning: A five years systematic review. *Journal of Computer Assisted Learning*, 40(6), 3005–3029. <https://doi.org/10.1111/jcal.13052>
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382–385. <https://doi.org/10.1097/00006199-198611000-00017>
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1), 84–99. <https://doi.org/10.1037/1082-989X.4.1.84>
- Matcha, W., Uzir, N. A., Gašević, D., & Pardo, A. (2020). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, 13(2), 226–245. <https://doi.org/10.1109/tlt.2019.2916802>
- Michael, J., Santoso, H. B., Junus, K., & Sandiwarno, S. (2025). Systematic review of self-regulated learning intervention and measurement in higher education: Towards a holistic, integrated, and technology-assisted model. *Journal of Information Technology Education: Research*, 24, Article 33. <https://doi.org/10.28945/5636>
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–173). Academic Press. [https://doi.org/10.1016/S0079-7421\(08\)60053-5](https://doi.org/10.1016/S0079-7421(08)60053-5)
- Newman, R. S. (1994). Adaptive help seeking: A strategy of self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and educational applications* (pp. 283–301). Lawrence Erlbaum Associates.
- Ngo, T. T. N., Chen, H. H. J., & Lai, K. K. W. (2024). The effectiveness of automated writing evaluation in EFL/ESL writing: A three-level meta-analysis. *Interactive Learning Environments*, 32(2), 727–744. <https://doi.org/10.1080/10494820.2022.2096642>
- Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading & Writing Quarterly*, 19(2), 139–158. <https://doi.org/10.1080/10573560308222>
- Paulsen, L., & Lindsay, E. (2024). Learning analytics dashboards are increasingly becoming about learning and not just analytics – A systematic review. *Education and Information Technologies*, 29(11), 14279–14308. <https://doi.org/10.1007/s10639-023-12401-4>
- Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). Academic Press. <https://doi.org/10.1016/B978-012109890-2/50043-3>
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Ponce-Renova, H. F. (2022). Comparing effect sizes and their confidence intervals: A primer on equivalence testing in educational research. *Journal of New Approaches in Educational Research*, 11(2), 209–225.

<https://doi.org/10.7821/naer.2022.7.930>

- Puntambekar, S., & Hübscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist*, 40(1), 1–12. [https://doi.org/10.1207/s15326985ep4001\\_1](https://doi.org/10.1207/s15326985ep4001_1)
- Richey, R. C., & Klein, J. D. (2007). *Design and development research: Methods, strategies, and issues*. Lawrence Erlbaum Associates.
- Sharma, K., Nguyen, A., & Hong, Y. (2024). Self-regulation and shared regulation in collaborative learning in adaptive digital learning environments: A systematic review of empirical studies. *British Journal of Educational Technology*, 55(4), 1398–1436. <https://doi.org/10.1111/bjet.13459>
- Shen, B., & Bai, B. (2024). Enhancing Chinese university students' writing performance and self-regulated learning (SRL) writing strategy use through a strategy-based intervention. *System*, 122, Article 103249. <https://doi.org/10.1016/j.system.2024.103249>
- Shi, H., & Aryadoust, V. (2024). A systematic review of AI-based automated written feedback research. *ReCALL*, 36(2), 187–209. <https://doi.org/10.1017/s0958344023000265>
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, 1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>
- Song, X., Razali, A. B., Sulaiman, T., & Jeyaraj, J. J. (2025). Effectiveness of online project-based learning on Chinese EFL learners' critical thinking skills and reading comprehension ability. *Thinking Skills and Creativity*, 56, Article 101778. <https://doi.org/10.1016/j.tsc.2025.101778>
- Su, L., Noordin, N., & Yang, D. (2023). Unraveling the path to success: Exploring self-regulated language learning among Chinese college EFL learners. *SAGE Open*, 13(4). <https://doi.org/10.1177/21582440231218537>
- Teng, L. S., & Zhang, L. J. (2020). Empowering learners in the second/foreign language classroom: Can self-regulated learning strategies-based writing instruction make a difference? *Journal of Second Language Writing*, 48, Article 100701. <https://doi.org/10.1016/j.jslw.2019.100701>
- Teng, M. F. (2025). Metacognitive awareness and EFL learners' perceptions and experiences in utilising ChatGPT for writing feedback. *European Journal of Education*, 60(1), Article e12811. <https://doi.org/10.1111/ejed.12811>
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249–276. <https://doi.org/10.3102/00346543068003249>
- Toulmin, S. E. (1958). *The uses of argument*. Cambridge University Press.
- van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review*, 22(3), 271–296. <https://doi.org/10.1007/s10648-010-9127-6>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wang, J., Graham, S., Kim, Y.-S. G., & Steiss, J. (2026). Zooming into two measurement issues in writing self-efficacy: Revision as a distinct dimension and the generality hypothesis in argumentative writing. *Reading and Writing*, 39(5), 1849–1876. <https://doi.org/10.1007/s11145-025-10679-z>
- Wang, X., Ma, J., Li, X., & Shen, X. (2023). Validation of self-regulated writing strategies for advanced EFL learners in China: A structural equation modeling analysis. *European Journal of Investigation in Health, Psychology and Education*, 13(4), 776–795. <https://doi.org/10.3390/ejihpe13040059>
- Watkins, D. A., & Biggs, J. B. (Eds.). (1996). *The Chinese learner: Cultural, psychological and contextual influences*. Comparative Education Research Centre; Australian Council for Educational Research.
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Lawrence Erlbaum Associates.
- Won, S., & Chang, Y. (2024). Antecedents and consequences of academic help-seeking in online STEM learning. *Frontiers in Psychology*, 15, 1438299. <https://doi.org/10.3389/fpsyg.2024.1438299>
- Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G.-J., & Paas, F. (2019). Supporting self-regulated learning in online learning environments and MOOCs: A systematic review. *International Journal of Human–Computer Interaction*, 35(4–5), 356–373. <https://doi.org/10.1080/10447318.2018.1543084>
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and*

- Psychiatry*, 17(2), 89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- Xu, W., & Zhu, X. (2024). Examining metacognitive strategy use in L1 and L2 task-situated writing: Effects, transferability, and cross-language facilitation. *Metacognition and Learning*, 19(3), 773–792. <https://doi.org/10.1007/s11409-024-09387-w>
- Yu, S., & Lee, I. (2016). Peer feedback in second language writing (2005–2014). *Language Teaching*, 49(4), 461–493. <https://doi.org/10.1017/S0261444816000161>
- Yun, H., Song, H.-D., & Kim, Y. (2025). Identifying university students' online self-regulated learning profiles: Predictors, outcomes, and differentiated instructional strategies. *European Journal of Psychology of Education*, 40(1), Article 5. <https://doi.org/10.1007/s10212-024-00907-5>
- Zhang, J., & Zhang, L. J. (2024). Profiling L2 students' writing self-efficacy and its relationship with their writing strategies for self-regulated learning. *System*, 122, 103253. <https://doi.org/10.1016/j.system.2024.103253>
- Zhang, J., Zhang, L. J., & Zhu, Y. (2023). Development and validation of a genre-based second language (L2) writing self-efficacy scale. *Frontiers in Psychology*, 14, 1181196. <https://doi.org/10.3389/fpsyg.2023.1181196>
- Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: A meta-analysis. *Asia Pacific Education Review*, 17(2), 187–202. <https://doi.org/10.1007/s12564-016-9426-9>
- Zhou, X., Chai, C.-S., Jong, M. S.-Y., & Feng, H. (2025). Needs satisfaction and online self-regulated learning among Chinese undergraduates. *PLOS ONE*, 20(4), Article e0321781. <https://doi.org/10.1371/journal.pone.0321781>
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press. <https://doi.org/10.1016/B978-012109890-2/50031-7>
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70. [https://doi.org/10.1207/s15430421tip4102\\_2](https://doi.org/10.1207/s15430421tip4102_2)