

Cultivating Online Civic Character through Deeper Learning: A Transformative Model for Digital Citizenship Education

Warlim Warlim

warlim_isya@upi.edu

Universitas Pendidikan Indonesia

Mupid Hidayat

Universitas Pendidikan Indonesia

Ratna Fitria

Universitas Pendidikan Indonesia

Aang Supriatna

Universitas Pendidikan Indonesia

Muhammad Imam Fahrizal

Universitas Pendidikan Indonesia

Research Article

Keywords: online civic character, digital citizenship education, deeper learning, transformative model, civic online reasoning, higher education

Posted Date: October 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11013912/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

Background. Digital citizenship education in higher education remains procedural and safety-oriented, regulating online conduct without cultivating the epistemic and moral dispositions democratic participation now requires; Global South evidence on deeper-learning alternatives is scarce.

Purpose. This study proposes and tests a transformative, deeper-learning-based model of online civic character education—integrating civic online reasoning, moral self-concept and civic purpose—over twelve weeks with undergraduates at an Indonesian teacher-education university.

Methods. A cluster-randomised, sequential explanatory mixed-methods design assigned twelve intact classes ($N = 342$, six faculties) to intervention or comparison conditions, measuring outcomes at pretest, posttest and eight-week follow-up with four instruments (DCS, CORA, MSC-D, CPS) with measurement invariance formally tested on the sample and analysed with multilevel mixed-effects models. The qualitative phase drew on observation, interviews (36 students, 6 lecturers) and student portfolios, analysed through reflexive thematic analysis.

Results. Significant condition $\times$ time effects emerged for three instruments, tracing a coherent gradient. Civic purpose showed the largest, most durable gain ($\eta^2p = .033$), followed by digital citizenship ($\eta^2p = .024$); moral self-concept improved only immediately, while the sole performance measure—civic online reasoning—showed a favourable non-significant trend that qualitative evidence reinterprets as a genuinely emerging skill. Effects were equitable across sex, socioeconomic status and digital access; thematic analysis identified contrasting pathways of transformation and stagnation.

Conclusion. Deeper learning differentially strengthened online civic character: its facets proved empirically separable, developing at different rates and durabilities, so must be cultivated—and measured—in their own right. Civic character can be cultivated pedagogically, not merely regulated, in digitally saturated environments.

1 Introduction

The digital age has fundamentally reshaped what it means to be a citizen: public participation now unfolds not only in physical space but also in fluid and complex online arenas [1–3]. This transformation calls for a redefinition of civic competence that encompasses the capacity to reason critically about information, to interact ethically in networked spaces, and to understand one's digital rights and responsibilities [4, 5]. Yet many current approaches to civic education remain procedural and do not reach the deeper dimensions of character needed to confront misinformation, polarisation and online hate speech [6, 7]. This challenge is increasingly urgent because young people spend a substantial share of their lives online while frequently lacking the reflective and moral capacities such environments demand [8, 9]. A transformative model is therefore required that systematically integrates deeper learning into online civic character education, so as to form digital citizens who are principled, critical and participatory [10, 11].

A central weakness of current digital citizenship education is the dominance of technical and normative framings that emphasise online safety and polite conduct while neglecting the development of civic reasoning and deliberative ethics [12–14]. Studies by Breakstone et al. [15] and Wineburg et al. [16] show that most secondary and tertiary students cannot reliably distinguish sponsored content from legitimate reporting, and fail to perform lateral reading — a foundational skill of digital citizenship. Without a character foundation such as intellectual honesty, open-mindedness and social responsibility, technical skill alone can be turned to spreading disinformation or reinforcing echo chambers [17, 18]. This underscores the urgency of integrating deeper learning — critical thinking, collaboration and moral-identity formation — into the digital citizenship curriculum [19, 20]. Online character education can thus no longer be separated from the strengthening of young people's epistemic and affective competencies [21, 22].

The deeper-learning dimension of digital citizenship demands a shift from mere information literacy toward the internalisation of deliberative-democratic values in everyday online practice [23, 24]. Choi and Cristol [4] emphasise the importance of intersectionality in understanding how social, economic and cultural background shapes how individuals experience and enact digital citizenship. Harris and Johns [25], meanwhile, show that approaches centred on individual resilience often fail to capture the collective and relational dimensions of online civic engagement. What is needed is a model that integrates civic reasoning in the tradition of McGrew [26], the moral self-concept described by Kim et al. [10], and authentic assessment of character of the kind developed by Winarno et al. [27]. Empirical work by Warlim et al. [11] indicates that using artificial intelligence as a reflective assistant in online character learning can strengthen the consistency of students' ethical conduct, provided it does not replace human reflection. The transformative model proposed here therefore emphasises challenging project-based learning, cross-perspective dialogue, and formative feedback on the development of online civic character [28, 29].

The principal contribution of this article is a model of intentional digital citizenship education that explicitly pairs deeper learning with the formation of online civic character – a synthesis still rare in the international literature [30, 31]. The meta-analysis by Lu and Gu [32] confirms that most digital citizenship studies remain fragmented between technical skill and participatory attitude, without addressing the character foundation that binds the two. Costa and Oliver [33] and Örtengren [9] show that the co-production of digital knowledge between teachers and students can be an effective medium for developing collective responsibility and ethical sensitivity. Using Rasch modelling, Prasetyo et al. [34] identify a gap between students' stated beliefs and their actual conduct in digital-citizenship contexts, indicating a need for structured character habituation. This article addresses that gap by integrating moral self-concept [10], civic purpose [35] and online reasoning [36] into a single integrative learning framework that can be implemented across educational levels.

Substantively, the model we propose rests on three pillars: (1) epistemic scaffolding to train online civic reasoning; (2) dialogic deliberation to develop empathy and tolerance; and (3) reflective digital action to build habits of ethical, responsible participation [37–39]. The first pillar adapts findings by McGrew and Breakstone [17] on how thematic digital-literacy lessons can significantly improve students' ability to evaluate contested claims. The second draws on Vallès-Peris and Domènech [23], who show that a pragmatic-democratic approach in online classrooms can cultivate a climate of mutual listening. The third develops from Harrison and Polizzi [40] on adolescents' online moral decision-making, and from Cleofas and Labayo [41], who link global competence to responsible online participation. Together these pillars produce a learning ecosystem that not only transfers knowledge but also forms character dispositions such as intellectual courage, ethical curiosity and commitment to the common good [42, 43]. Digital citizenship education thereby moves beyond *learning about* toward *becoming* a mature citizen in the digital age [44, 45].

This article sets out, systematically, the theoretical framework and implementation of a deeper-learning-based online civic character education model, and reports evidence from a bounded twelve-week trial of its effectiveness in strengthening students' civic reasoning and moral resilience. The model is intended as a reference for curriculum developers, educators and researchers designing interventions that are not merely pragmatic but oriented toward forming democratic, principled digital citizens [46–48]. Framing civic education in the digital age in this way asserts our theoretical position: it must be willing to move beyond technical literacy toward sustained character transformation [49, 50]. This is where the novelty and urgency of the proposed model lie, and it provides the foundation for the sections that follow.

2 Theoretical framework

2.1 Philosophical foundations of digital citizenship

Digital transformation has changed the nature of citizenship from a legal status into a fluid, participatory practice that crosses spaces [1–3]. Westheimer and Kahne [37] distinguish three types of citizen – personally responsible, participatory and justice-oriented – the last of which stresses the capacity to analyse the structural roots of injustice. In the online domain, Choi [13] defines digital citizenship as the ability to use the internet ethically, critically and participatorily to build a democratic society. Many curricula, however, remain confined to the personally responsible dimension, such as online safety, without developing critical awareness [5, 9]. Yet Kahne and Bowyer [6] show that young people exposed to polarised political content encounter serious difficulties with motivated reasoning. A philosophical foundation is therefore needed that integrates democratic pragmatism [23] and an ethic of digital care [45] to form a transformative framework for online citizenship [19, 24].

2.2 Deeper learning as the core of online character education

Deeper learning is defined as the development of higher-order cognitive competencies, affective dispositions and collaborative skills that enable individuals to transfer knowledge to new situations [47, 51]. In the digital-citizenship context this approach becomes crucial because online challenges are non-routine and laden with moral dilemmas [10, 40]. Kim et al. [10] find that moral self-concept acts as a catalyst linking civic competence to civic purpose, especially when students are engaged in digital-dilemma simulations. McGrew [7] shows that a civic-online-reasoning intervention designed on deeper-learning principles significantly improves students' ability to evaluate contested claims. The project-based model developed by Prasetyo et al. [28] likewise demonstrates that pre-service teachers engaged in online civic projects show gains in ethical conduct and digital responsibility [35]. Deeper learning is thus a pedagogical foundation inseparable from the cultivation of online civic character [21, 46].

2.3 Civic online reasoning

Civic online reasoning (COR), developed by the Stanford History Education Group, is an epistemic framework for evaluating the credibility of digital information, particularly on matters of public concern [16, 36]. COR comprises three core practices: lateral reading (verifying claims by opening new tabs), click restraint (delaying the click on the first result), and the identification of hidden bias and agendas [15, 26]. A national study by Breakstone et al. [15] found that fewer than one in five students could perform lateral reading competently, with the majority swayed by a site's visual appearance. An intervention by McGrew and Breakstone [17] showed that COR lessons integrated into a social-studies curriculum over six weeks produced substantial gains in source-evaluation ability. Yet COR alone is insufficient without supporting character dispositions such as healthy scepticism and epistemic humility [50, 52]. In our framework, therefore, COR is positioned as the cognitive dimension that must be governed by ethics and character [18, 29].

2.4 Civic character in moral and affective perspective

Online civic character education cannot be reduced to technical skill; it concerns the formation of moral emotions and civic virtues such as digital empathy, the courage to report disinformation, and collective responsibility [19, 45]. The moral-decision theory of Harrison and Polizzi [40] explains that adolescents face three kinds of online dilemma: the civility dilemma (whether to respond to an offensive comment), the accuracy dilemma (whether to share unverified content), and the solidarity dilemma (whether to defend a peer who is being cyberbullied). Cleofas and Labayo [41] show that global competence – including respect for cross-cultural perspectives – correlates positively with responsible digital-citizenship practice. Kim et al. [10] further find that a strong moral self-concept can shield against the negative effects of algorithmic echo chambers. An integrative model of character education thus cannot ignore the affective and relational dimensions [23, 25, 48].

2.5 Democratic participation and online civic engagement

Transformative digital citizenship must be oriented toward strengthening democratic participation, not mere compliance with online rules [4, 38]. Bennett et al. [1] distinguish actualising citizenship, which centres on participatory self-expression, from the more traditional, loyalty-based dutiful citizenship. Kahne and Bowyer [53] demonstrate that digital media-literacy education emphasising the analysis of contemporary issues can meaningfully increase adolescents' online political engagement. Mirra et al. [54], meanwhile, criticise digital-citizenship approaches that are too narrow because they ignore tough issues such as algorithmic racism and access inequality. A more equitable approach [4] uses an intersectional lens to understand how overlapping social identities shape the experience of online citizenship. In our framework, democratic participation is measured not only by the frequency of online activity but also by the quality of deliberation, solidarity with marginalised groups, and commitment to factual truth [8, 43, 55].

2.6 The role of technology and artificial intelligence in character formation

Technological advances, especially generative AI, carry twofold implications for civic character education [11, 44]. On one hand, AI can act as a reflective assistant that helps students identify cognitive biases and the ethical consequences of their online actions [18]. Warlim et al. [11] show that students using an AI-enhanced feedback system in online discussion display greater consistency in ethical conduct and greater awareness of others' digital rights. On the other hand, automation bias and the erosion of moral autonomy become serious threats where technology replaces human reflection [19, 49]. The use of AI in character education must therefore be augmentative, not substitutive. Valdez et al. [22] additionally find that strong digital citizenship can prevent online cheating and promote academic delay of gratification – indicating that character formation matters not only for democratic participation but also for academic integrity [9, 14, 42].

2.7 An integrative model: character, reasoning and participation

On the basis of the synthesis above, we propose a model of transformative digital citizenship character education that integrates three domains: (1) epistemic virtues (intellectual honesty, healthy scepticism, lateral reading); (2) moral emotions (digital empathy, moral courage and concern for others' rights); and (3) civic actions (deliberative participation, advocacy and collaboration across difference) [13, 37, 38]. The framework adopts deeper learning as the pedagogical approach that bridges these three domains through authentic projects, ethical reflection and deliberative dialogue [28, 47]. Costa and Oliver [33] confirm that co-producing digital-literacy practices with students builds a sense of ownership and collective responsibility, while the longitudinal work of Kim et al. [10] shows that structured digital civic learning can strengthen civic purpose in early adolescence [35, 56]. The model is cyclical rather than linear: each online action becomes material for ethical evaluation that informs subsequent learning [29, 30, 40]. In operational terms, the three pedagogical pillars map onto the study's four measured constructs as a one-to-many correspondence: epistemic scaffolding feeds civic online reasoning (CORA); dialogic deliberation feeds moral self-concept (MSC-D); and reflective digital action feeds both digital citizenship (DCS) and civic purpose (CPS), as set out in the joint display (Table 10).

2.8 Implications for research and measurement

Operationalising this model requires valid and reliable measurement, given the complexity of the online-civic-character construct [2, 57]. Choi et al. [2] developed the Digital Citizenship Scale with five dimensions: internet political activism, technical skills, local/global awareness, critical perspective and networking/interpersonal communication. That scale, however, does not explicitly measure character dimensions such as digital empathy and moral courage. Recent work by Kim et al. [10] and Warlim et al. [11] seeks to integrate moral self-concept and AI-facilitated ethical reflection into measurement. In the Indonesian context, Prasetyo et al. [34] used Rasch analysis to validate a digital-citizenship instrument, revealing an item–person map that signals the need for more intensive character education. Measurement

must also account for contextual factors such as socioeconomic status and technology access [4, 32]. Consistent with the design reported below, we treat online civic character as a multidimensional construct: civic online reasoning (CORA), moral self-concept (MSC-D) and civic purpose (CPS) are modelled as distinct but related constructs rather than as a single scale, and the possibility of a higher-order structure is examined empirically (see Section 4.1) [31, 33]. This framework grounds the intervention and the hypotheses to be tested in the subsequent sections [22, 58].

3 Methods

3.1 Design and approach

This study used a sequential explanatory mixed-methods design [59]. The quantitative phase employed a cluster-randomised pretest–posttest–follow-up comparison-group design: because the deeper-learning intervention is delivered to whole classes, intact classes—not individual students—were the unit of random assignment, avoiding within-class contamination between conditions. The qualitative phase used a multi-site case study (participant observation, interviews and documented ethical reasoning) to examine the mechanisms of character change in context. Methodological triangulation is warranted because online civic character is multidimensional and difficult to capture with a single method [14, 31]. The population comprised second-semester students at Universitas Pendidikan Indonesia (UPI) with varying levels of digital access.

3.2 Participants and sampling

Participants were 342 students from six UPI faculties. Sampling proceeded in two stages. First, within each faculty, intact second-semester classes taught by Civic Education (PKN) lecturers took part voluntarily; internet access, data-plan support and participation were recorded descriptively for the sample rather than used to screen individuals. Second, because the intervention runs per class over twelve weekly sessions, two intact classes per faculty were randomly assigned at the class level (cluster randomisation) using a random-number table stratified by faculty—one to the intervention and one to the comparison condition—yielding twelve classes (six per arm; 172 intervention, 170 comparison students). The intervention arm received the deeper-learning-based online civic character education model over twelve weeks; the comparison arm received conventional digital-citizenship instruction. Because assignment was at the class level, the data were analysed with multilevel mixed-effects models reflecting the nested structure (students within classes within faculties), with class as a random effect (§ 3.6). At the sample level, participants were characterised by active daily social-media use (at least two hours per day), independent device and internet access, and no prior training in civic online reasoning.

Statistical power was evaluated at the design level. With six clusters per arm (~ 28–29 students each; $N = 342$), a two-sided α of .05, and a conservative intraclass correlation of $ICC \leq 0.05$ (design effect ≈ 2.4), the design had approximately .80 power to detect a between-arm standardised difference of $d \approx 0.50$. Two features raised the realised sensitivity: the primary estimand is the baseline-adjusted condition $\times$ time interaction—more efficient than a posttest-only contrast—and the empirically negligible between-cluster variance ($ICC \approx 0$; § 4.1) reduced the realised design effect to ≈ 1 , giving a minimum detectable effect near $d = 0.30$. The study was therefore sensitive to the small-to-medium effects observed for most constructs (posttest Hedges $g \approx 0.32$ – 0.36) but underpowered for the smallest effect (civic online reasoning, $g \approx 0.12$), which is accordingly interpreted as a directional trend rather than a confirmed null (§ 4.2.2, § 5.7). With only twelve clusters, inference about the class-level condition effect rests on limited between-cluster degrees of freedom; the near-zero ICC and the convergence of the multilevel, pretest-adjusted (ANCOVA) and non-parametric (Mann–Whitney) analyses mitigate this constraint.

3.3 Quantitative instruments

Four instruments, all adapted into Indonesian by back-translation, were used. The Digital Citizenship Scale (DCS) [2, 34] comprises 25 items across five dimensions (internet political activism, technical skills, local/global awareness, critical perspective, networking/interpersonal communication). The Civic Online Reasoning Assessment (CORA), built on the civic-online-reasoning framework [15, 17] and operationalised as a Pancasila-grounded, forced-choice scenario test, comprises eight items aggregated into a 0–11 per-occasion total: items 1, 2, 6 and 8 are two-tier (scored 1 only when both parts are correct); items 3, 4 and 7 are single best-choice (0/1); and item 5 is select-all, scored as the number of correct minus incorrect options and bounded 0–4. Three parallel forms of equivalent difficulty were administered by occasion (Form A at pretest, B at posttest, C at follow-up), their cross-occasion equivalence supported by the full longitudinal scalar invariance reported in § 4.1 (Table 4); this forced-choice format replaced an earlier task-based version that had proved too burdensome and yielded incomplete responses. The Moral Self-Concept Scale in Digital Context (MSC-D) [10, 40] comprises 20 items across four dimensions (digital honesty, digital empathy, digital responsibility, digital moral courage). The Civic Purpose Scale (CPS) [10, 35] comprises 8 items and, consistent with the theoretical model (§ 2.7), is treated as a construct distinct from moral self-concept.

The reliability and validity values reported in the source studies are used only as an initial indication of content adequacy, not as this study's psychometric values. All instruments were first piloted with 60 students outside the research sample; a data-screening step (detection of zero-variance items, inspection of the inter-item correlation matrix for singularity, and inspection of response patterns for non-genuine responding) was conducted, and items with factor loadings < 0.50, AVE < 0.50, or reliability below threshold ($\alpha/\omega < 0.70$) were revised before pretest data collection. The measurement properties on which all inferential analyses rest—confirmatory factor structure and fit, reliability (Cronbach's α and McDonald's ω), convergent validity (AVE), discriminant validity (Fornell–Larcker criterion and the heterotrait–monotrait ratio, HTMT), the dimensionality of the online-civic-character construct (a second-order confirmatory factor analysis), and longitudinal and group measurement invariance—were then computed on the analytic sample ($N = 342$) and are reported at the opening of the Results (§ 4.1, Tables 1–5). Because CORA is scored objectively against a fixed key, no inter-rater statistic applies to it; inter-rater agreement is instead reported for the observational behaviour codes (§ 3.5) and the thematic coding (§ 3.7). The pretest was administered one week before the intervention, the posttest one week after it ended, and a follow-up eight weeks later to assess durability; testing conditions (device and setting) were standardised across participants at each occasion.

3.4 The deeper-learning intervention

The intervention rested on four deeper-learning elements: (1) meaningful, socially relevant content; (2) critical thinking and complex problem-solving; (3) effective collaboration and communication; and (4) self-directed learning with sustained ethical reflection [47, 51]. Across twelve weekly 90-minute sessions, students engaged in digital-dilemma simulations, lateral reading on contested issues, cross-perspective deliberative discussion, and data-based social-advocacy projects [17, 40]. Each session opened with a case-based ethical dilemma and proceeded to small-group inquiry using a civic-online-reasoning protocol [16, 23]; after every third session (Sessions 3, 6, 9 and 12) students compiled a digital-citizenship portfolio documenting their reasoning and moral reflection [33, 45]. The comparison group received conventional instruction on online safety and digital ethics without a deeper-learning approach [14, 58].

Implementation fidelity was monitored in all twelve intervention sessions using a structured checklist (Fidelity of Implementation Checklist, Supplementary Instrument), and comparison-group sessions were checked with a binary differentiation checklist to confirm the absence of deeper-learning elements. This continuous session-by-session monitoring is distinct from the three-point participant observation described in § 3.5, which targets the emergent dynamics of collaborative moral reasoning at selected junctures; the three-point differentiation check in Appendix B.1 is a time-anchored subset of the continuous monitoring, aligned to the measurement schedule for the joint-display analysis (§ 3.8).

3.5 Qualitative data collection

Qualitative data were gathered through three complementary channels, each capturing a different temporal grain of character development.

Participant observation, using a protocol adapted from Black et al. [8], recorded the dynamics of collaborative moral decision-making at three purposively selected sessions—Session 1 (coinciding with the pretest), Session 6 (the midpoint) and Session 12 (coinciding with the posttest)—so that the observational record aligns with the quantitative measurement schedule and supports joint-display comparison at matched time-points (§ 3.8). The three points capture, respectively, a baseline record of classroom reasoning dynamics, an interim indicator of change, and the cumulative behavioural pattern at posttest. In the comparison classes, observation at the corresponding three sessions was restricted to a treatment-differentiation check and a general engagement indicator; the full behaviour-coding scheme (lateral reading, click restraint, trust shift, expressed empathy, moral courage, cognitive dilemma, reflexive correction, collaborative behaviour; Appendix B.1) was applied only to the intervention qualitative subsample ($n = 18$) whose participant codes are shared with the interview and portfolio data—not to the full intervention cohort ($n = 172$) or to the comparison classes—because those codes operationalise deeper-learning activities that are absent by design from conventional instruction. At least 20% of observation sessions at each time-point were double-coded independently by two trained observers, with agreement computed as Cohen's κ for the behaviour codes and as percentage agreement for the binary checklist (reported in § 4). Because direct observation was limited to three points, continuity across the intervening weeks was maintained through the portfolios—compiled after every third session—and through the retrospective accounts obtained in the post-intervention interviews.

Semi-structured interviews were conducted with 36 students (18 intervention, 18 comparison) and 6 lecturers representing each faculty, using two guides: an intervention-group guide referencing the specific learning activities and a comparison-group guide worded so as not to presuppose participation in them. Documented ethical reasoning was drawn from the posts, comments and lateral-reading processes recorded in the digital-citizenship portfolios, with participant anonymity maintained throughout. All interviews were audio-recorded, transcribed verbatim and returned to participants for member checking. Interviewing continued until thematic saturation, assessed primarily from the interview and portfolio corpus given the intentionally bounded observational data.

3.6 Quantitative analysis

Analysis followed an intention-to-treat approach using SPSS 29 and R 4.3 [62]. Because the unit of random assignment was the class (§ 3.2), the primary analysis used multilevel mixed-effects models: each instrument's scores were regressed on condition, time (pretest/posttest/follow-up) and the condition $\times$ time interaction as fixed effects, with class as a random intercept controlling for the intraclass dependence of students within the same class. Parametric assumptions were examined beforehand (Kolmogorov–Smirnov test of normality; Levene's test of homogeneity), and independent-samples t-tests, Mann–Whitney U tests and pretest-adjusted ANCOVA served as convergent, supplementary analyses. Effect sizes were computed as Cohen's d/dz , Hedges g and partial eta-squared (η^2p) and interpreted by Cohen's conventions [60] (small ≥ 0.20 , medium ≥ 0.50 , large ≥ 0.80). Moderator analysis added condition $\times$ moderator interactions (sex, socioeconomic status, digital access) within the same model, with the resulting tests adjusted by the Benjamini–Hochberg false-discovery-rate procedure. Implementation fidelity was recorded per session and examined descriptively at the class level; because between-cluster degrees of freedom are limited (twelve clusters), formal covariate adjustment and an adequately powered Condition $\times$ Fidelity moderation test lie beyond the present design and are identified as priorities for future work (§ 5.7).

3.7 Qualitative analysis

Interview, observation and portfolio data were analysed using Braun and Clarke's [63] six-phase reflexive thematic analysis—familiarisation, inductive descriptive and in-vivo coding, searching for and reviewing candidate themes, defining and naming themes, and writing up—as applied to digital-citizenship contexts by Richardson et al. [14] and Fajri et al. [30]. Candidate themes included epistemic shifts, the negotiation of moral emotions, and the emergence of civic agency, with operational definitions agreed among the researchers and each theme supported by representative quotations. Analysis was supported by NVivo 14, and inter-coder reliability (Cohen's κ) is reported in the Results.

3.8 Data integration and trustworthiness

The two strands were integrated through joint-display analysis and an iterative weaving approach [64]. Meta-inferences were drawn by comparing each statistical result with the corresponding qualitative narrative to confirm or enrich it; because the three observation points are anchored to the pretest, midpoint and posttest occasions, the joint display aligns observational excerpts with the quantitative score at each matching occasion, strengthening interpretability. Trustworthiness was strengthened through prolonged engagement, triangulation of sources and methods, and a documented audit trail [61]; transferability through thick description of context and participants; dependability through a code–recode procedure; and confirmability through peer debriefing with three senior researchers in civic education and digital literacy. The bounded three-point observation schedule is a recognised limitation—partly offset by the denser portfolio record and the retrospective interviews and revisited in § 5.7. The study received ethics approval from the university ethics committee, and informed consent was obtained from all participants.

4 Results

The results integrate the quantitative and qualitative strands of the sequential explanatory design (§ 3.1, § 3.8): § 4.1–4.4 establish the measurement model, invariance and baseline equivalence and report each construct's effect and its equity across subgroups; § 4.5–4.7 present the qualitative, observational and fidelity evidence; and § 4.8–4.10 integrate the strands and specify boundary conditions. Differential effects are indexed by the condition $\times$ time interaction (§ 3.6); effect sizes follow Cohen's conventions (small ≥ 0.20 , medium ≥ 0.50 , large ≥ 0.80).

4.1 Preliminary quantitative analyses

Complete data at pretest, posttest and eight-week follow-up were available for all 342 randomised students (intervention $n = 172$; comparison $n = 170$; no missing responses). Because assignment was at the class level, every model carried a random intercept for the clustering unit; between-cluster variance was negligible for all outcomes (ICC ≈ 0), so multilevel and single-level estimates coincided.

Measurement model, reliability and validity. Confirmatory factor analysis supported each instrument's a priori structure (Table 1): the five-factor DCS (CFI = .977, RMSEA = .025), four-factor MSC-D (CFI = .980, RMSEA = .029) and one-factor CPS (CFI ≈ 1.00 , RMSEA = .000) all fitted well, with standardised loadings $\geq .59$ (CPS $\geq .52$). Internal consistency was good (Table 2), and where average variance extracted fell below .50 (AVE = .39–.56) composite reliabilities of .76–.86 preserved convergent validity under the Fornell–Larcker/Malhotra criterion [66, 67]. CORA was acceptably reliable ($\alpha = .67$, $\omega = .79$). Discriminant validity was clear (HTMT = .13–.31; Table 3). A second-order single-factor model fitted poorly (CFI = .35, TLI = .33, RMSEA = .081) with a degenerate loading pattern (DCS = 1.00, MSC-D = .07, CPS = .02, CORA = .03). Given the low inter-construct HTMT values (.13–.31; Table 3), this degenerate fit was anticipated, and the second-order model was fitted to confirm the structure rather than in expectation of a higher-order factor. The four facets are therefore related but separable and are modelled as distinct constructs, the DCS and MSC-D at dimension level (§ 2.7).

Table 1
Measurement properties of the four instruments (analytic sample, N = 342).

Instrument	Model / dimensions	Items	CFI	TLI	RMSEA	Reliability (α / ω)	Convergent & discriminant validity
DCS	5-factor (IPA, TS, LGA, CP, NI)	25	.977	.975	.025	α .76–.84 / ω .84–.88	AVE .39–.50; loadings $\geq$.59; HTMT $\leq$.31
MSC-D	4-factor (DH, DE, DR, DMC)	20	.980	.977	.029	α .81–.86 / ω .87–.90	AVE .46–.56; loadings $\geq$.63; HTMT $\leq$.31
CPS	1-factor (civic purpose)	8	$\approx$ 1.00	1.00	.000	α .82 / ω .86	AVE .37; CR .82; loadings $\geq$.52
CORA †	1-factor (performance)	8 tasks	—	—	—	α .67 / ω .79	Single 0–11 reasoning score

Note. Values computed on pretest analytic-sample data (N = 342). Composite reliability (CR) > .70 offsets AVE < .50 (Fornell–Larcker/Malhotra criterion) [66, 67]. A second-order “online civic character” model over the four constructs fitted poorly (CFI = .35, TLI = .33, RMSEA = .081), confirming a distinct-but-related structure. † CORA is a forced-choice scenario test of eight items per form aggregated into a single 0–11 total (seven dichotomous items plus one 0–4 polytomous item); being key-scored, it carries no inter-rater κ /ICC. IPA = internet political activism; TS = technical skills; LGA = local/global awareness; CP = critical perspective; NI = networking/interpersonal; DH = digital honesty; DE = digital empathy; DR = digital responsibility; DMC = digital moral courage.

Table 2
Subscale-level reliability and convergent validity (analytic sample, N = 342).

Instrument	Subscale	k	α	ω	CR	AVE	$\sqrt{\text{AVE}}$	Min. loading
DCS	IPA — Internet political activism	5	.835	.883	.835	.504	.710	.676
DCS	TS — Technical skills	5	.830	.880	.830	.495	.704	.677
DCS	LGA — Local/global awareness	5	.832	.882	.833	.499	.706	.679
DCS	CP — Critical perspective	5	.806	.866	.807	.457	.676	.624
DCS	NI — Networking/interpersonal	5	.762	.840	.762	.391	.626	.597
MSC-D	DH — Digital honesty	5	.823	.876	.823	.483	.695	.671
MSC-D	DE — Digital empathy	5	.807	.866	.808	.457	.676	.631
MSC-D	DR — Digital responsibility	5	.837	.885	.838	.509	.713	.661
MSC-D	DMC — Digital moral courage	5	.861	.900	.862	.557	.746	.680
CPS	Civic purpose	8	.820	.864	.822	.370	.609	.519

Note. k = number of items; CR = composite reliability; AVE = average variance extracted; $\sqrt{\text{AVE}}$ reported for the Fornell–Larcker comparison. CORA is scored as a single performance total and is not decomposed into subscales.

Table 3
Discriminant validity – heterotrait–monotrait (HTMT) ratios (analytic sample, N = 342).

DCS	DCS	MSC-D	CPS	CORA
	—	.31	.21	.25
MSC-D		—	.21	.24
CPS			—	.13
CORA				—
<i>Note. All HTMT ratios $\leq .31$, well below the conventional .85 threshold, indicating clear discriminant validity among the four constructs.</i>				

Measurement invariance. Because scores are compared across three occasions and two groups, invariance was tested before interpreting means (Tables 4, 5). Parcel-based ML models were judged by $\Delta CFI \leq .010$ and $\Delta RMSEA \leq .015$ [65]; for DCS and MSC-D, whose weakly correlated dimensions destabilise incremental indices, the χ^2 -difference test and $\Delta RMSEA$ governed. Across occasions all four instruments reached full scalar invariance. Between groups, configural invariance held for all four and metric invariance for DCS, MSC-D and CORA; CORA and MSC-D additionally reached scalar invariance, DCS met only partial scalar invariance ($\Delta\chi^2(4) = 10.36$, $p = .035$; $\Delta RMSEA = +.022$), and CPS showed loading non-invariance ($\Delta\chi^2(3) = 22.86$, $p < .001$; $\Delta RMSEA = +.086$).

Table 4
Longitudinal measurement invariance across pretest–posttest–follow-up (analytic sample, N = 342).

Instrument	Model	χ^2	df	CFI	RMSEA	$\Delta\chi^2$ (df)	p	ΔCFI	$\Delta RMSEA$
DCS	Configural	54.16	72	1.000	.000	—	—	—	—
	Metric	72.44	80	1.000	.000	18.27 (8)	.019	+ .000	+ .000
	Scalar	83.60	88	1.000	.000	11.16 (8)	.193	+ .000	+ .000
CORA	Configural	44.33	39	.996	.020	—	—	—	—
	Metric	54.34	45	.993	.025	10.01 (6)	.124	-.003	+ .005
	Scalar	58.54	51	.994	.021	4.21 (6)	.649	+ .001	-.004
MSC-D	Configural	40.65	39	1.000	.011	—	—	—	—
	Metric	54.42	45	.997	.025	13.77 (6)	.032	-.002	+ .014
	Scalar	63.87	51	.996	.027	9.46 (6)	.149	-.001	+ .002
CPS	Configural	47.14	39	.997	.025	—	—	—	—
	Metric	62.00	45	.994	.033	14.87 (6)	.021	-.003	+ .009
	Scalar	69.80	51	.994	.033	7.80 (6)	.253	-.001	-.000

Note. Parcel-based ML CFA, validated against an established SEM package at the configural level. Invariance criterion: $\Delta CFI \leq .010$ and $\Delta RMSEA \leq .015$ [65]. All four instruments reached full scalar invariance, so across-occasion mean comparisons are warranted.

Table 5

Measurement invariance between experimental and comparison groups at pretest (analytic sample, N = 342).

Instrument	Model	χ^2	df	CFI	RMSEA	$\Delta\chi^2$ (df)	p	Δ CFI	Δ RMSEA
DCS	Configural	5.80	10	1.000	.000	—	—	—	—
	Metric	10.56	14	1.000	.000	4.76 (4)	.313	+ .000	+ .000
	Scalar	20.92	18	.000 †	.022	10.36 (4)	.035	-1.000 †	+ .022
CORA	Configural	10.92	4	.972	.071	—	—	—	—
	Metric	11.57	7	.982	.044	0.65 (3)	.884	+ .009	-.027
	Scalar	12.41	10	.990	.027	0.84 (3)	.839	+ .009	-.017
MSC-D	Configural	7.93	4	.172 †	.054	—	—	—	—
	Metric	11.80	7	.000 †	.045	3.87 (3)	.276	-.172 †	-.009
	Scalar	13.84	10	.190 †	.034	2.05 (3)	.563	+ .190 †	-.011
CPS	Configural	2.01	4	1.000	.000	—	—	—	—
	Metric	24.87	7	.965	.086	22.86 (3)	< .001	-.035	+ .086
	Scalar	24.88	10	.971	.066	0.01 (3)	1.000	+ .006	-.020

Note. † For DCS and MSC-D the incremental fit indices (CFI, Δ CFI) are uninformative because their dimensions are only weakly intercorrelated (a near-fitting independence baseline makes CFI unstable); the χ^2 -difference test and Δ RMSEA govern the decision. Configural invariance held for all four instruments; metric invariance held for DCS, MSC-D and CORA; CORA and MSC-D additionally reached scalar invariance. DCS met metric but only partial scalar invariance ($\Delta\chi^2(4) = 10.36$, $p = .035$; Δ RMSEA = + .022). CPS showed loading (metric) non-invariance across groups ($\Delta\chi^2(3) = 22.86$, $p < .001$; Δ RMSEA = + .086).

Baseline equivalence and assumptions. At pretest the groups differed on no instrument or dimension (all $p > .05$; |Hedges g | ≤ 0.13), confirming equivalent groups and licensing the condition $\times$ time interaction as a treatment effect. Distributions were normal (non-significant Kolmogorov–Smirnov), Mann–Whitney U tests corroborated every conclusion below, and Greenhouse–Geisser corrections were applied where sphericity failed (DCS, CPS, MSC-D).

4.2 Effects of the intervention on online civic character

4.2.1 Digital citizenship (DCS)

The intervention produced a significant, differential gain in digital citizenship: the condition $\times$ time interaction was significant, $F(2, 680) = 8.33$, $p < .001$, $\eta^2p = .024$, with a medium within-group gain ($dz = 0.56$) and a posttest advantage over the comparison group ($g = 0.35$, $p = .001$). The within-group gain was retained at follow-up ($dz = 0.43$, $p < .001$), though the between-group advantage attenuated to a trend ($g = 0.18$, $p = .097$). The effect concentrated in the civic dimensions — internet political activism ($p = .006$, $\eta^2p = .015$) and local/global awareness ($p = .046$, $\eta^2p = .009$) — and

was absent for technical skills ($p = .48$), critical perspective ($p = .16$) and networking/interpersonal communication ($p = .11$).

4.2.2 Civic online reasoning (CORA)

For demonstrated civic online reasoning the effect was positive but non-significant. Both groups improved over time ($F(2, 680) = 16.74, p < .001$), and the intervention group's within-group gain ($dz = 0.30$) was about twice the comparison group's, yet the condition $\times$ time interaction was non-significant, $F(2, 680) = 1.84, p = .160, \eta^2p = .005$, with a small posttest difference ($g = 0.12, p = .25$). Convergent ANCOVA, mixed-model and gain-score analyses placed the posttest effect at $p \approx .07-.08$, and Mann-Whitney tests agreed.

4.2.3 Moral self-concept (MSC-D)

Moral self-concept showed a small, significant but transient effect: the interaction was significant, $F(2, 680) = 4.29, p = .014, \eta^2p = .013$, with a small-to-medium within-group gain ($dz = 0.43, p < .001$) and a posttest advantage ($g = 0.32, p = .004$) that disappeared by follow-up ($g = 0.14, p = .20$). No single dimension reached omnibus significance; in pretest-adjusted analyses only digital empathy was individually significant at posttest ($p = .034$), with digital honesty ($p = .055$) and digital responsibility ($p = .078$) borderline.

4.2.4 Civic purpose (CPS)

The clearest and most durable effect was for civic purpose: the interaction was the largest of the four, $F(2, 680) = 11.50, p < .001, \eta^2p = .033$, with a medium within-group gain ($dz = 0.60$) and a posttest advantage ($g = 0.36, p < .001$) that, though partly attenuated, remained significant at follow-up ($g = 0.26, p = .017$), whereas the comparison group's smaller initial gain disappeared. Pretest-adjusted and mixed-model analyses confirmed the advantage at both occasions (ANCOVA $p < .001$). As noted in § 4.1, a few CPS parcels showed baseline loading non-invariance; given pretest-equivalent means under class-level randomisation, the effect is read as robust in direction while acknowledging the measurement caveat.

4.3 Cross-instrument synthesis

Read together (Table 6; Fig. 5), the four outcomes trace a coherent gradient: effects were strongest and most durable for civic purpose (significant and retained at follow-up) and digital citizenship (significant, within-group gain retained, between-group advantage narrowing to a trend), intermediate and transient for moral self-concept (significant only immediately), and a non-significant positive trend for the single performance measure, civic online reasoning. All effects were small-to-moderate (posttest $g = 0.12-0.36$), as expected of a bounded twelve-week whole-class intervention.

Table 6
Cross-instrument summary of intervention effects (N = 342).

Instrument (outcome)	Type	C × T F(2,680)	p	η ² p	EXP dz pre→post	Between- g posttest	Between- g follow- up	Verdict
Civic purpose (CPS)	Self-report	11.50	< .001***	0.033	0.60	0.36***	0.26*	Significant and durable
Digital citizenship (DCS)	Self-report	8.33	< .001***	0.024	0.56	0.35**	0.18 ns	Significant; retained within-group, between-group trend at FU
Moral self-concept (MSC-D)	Self-report	4.29	0.014*	0.013	0.43	0.32**	0.14 ns	Significant but transient
Civic online reasoning (CORA)	Performance	1.84	0.160 ns	0.005	0.30	0.12 ns	0.09 ns	Non-significant positive trend

Note. Type = self-report disposition vs performance test. dz = within-group Cohen's dz (intervention group). Between-g = Hedges g (intervention – comparison) at each occasion. *** $p < .001$, ** $p < .01$, * $p < .05$, ns = not significant.

4.4 Moderator analysis: equity of effect

We tested whether effects varied by three pre-specified moderators — sex, socioeconomic status (SES) and digital access — each balanced across conditions (all $\chi^2 p \geq .44$). For every instrument the condition × moderator interaction on posttest and follow-up gains was tested, the 24 resulting tests adjusted by the Benjamini–Hochberg false-discovery-rate procedure. Effects were statistically uniform: five interactions reached the uncorrected threshold (raw $p < .05$) but none survived correction, and each explained under 2% of gain variance ($\eta^2 p < .02$; Fig. 6). The five raw signals were sex on CORA at posttest ($p = .018$) and follow-up ($p = .041$), sex on DCS at follow-up ($p = .030$), SES on CPS at follow-up ($p = .040$), and digital access on MSC-D at follow-up ($p = .030$). Each rested on fragile contrasts — notably the small male subsample ($n \approx 38$ per arm; e.g., DCS follow-up $g = 0.72$ confined to males) — and is reported as hypothesis-generating only. The programme thus benefited lower- and higher-resource students comparably rather than widening gaps.

4.5 The qualitative and observational evidence base

The statistical pattern records whether outcomes changed but not why the constructs responded so differently. The qualitative, observational and fidelity strand (§ 3.5, § 3.7) drew on four converging sources: semi-structured interviews with 36 students (18 intervention, 18 comparison) and 6 lecturers; participant observation of 18 intervention students across three of the six classes (Sessions 1, 6 and 12); ethical reasoning documented in portfolios; and a Fidelity-of-Implementation Checklist completed every session across all six classes (72 in total). Braun and Clarke's six-phase reflexive thematic analysis identified two contrasting causal chains: a transformation pathway in the intervention group (deeper-learning experience → active verification → reflective pause → empathetic correction → open accountability → acknowledged transformation) and a stagnation pathway in the comparison group (declarative

learning → passive verification → impulsive sharing → conflict avoidance → silent avoidance → acknowledged stagnation), elaborated through the seven themes of § 4.6.

4.5.1 Implementation fidelity as a precondition for interpreting effects

Implementation fidelity. Before attributing effects to the design, delivery quality was examined. Across the six classes the mean fidelity score on the 16-item core checklist (maximum 22) was 17.0 (77.3%), in the ‘moderate’ band and close to the ‘high’ threshold ($\geq 80\%$): one class (16.7%) was ‘high’ and five (83.3%) ‘moderate’, none ‘low’, so no class suffered a systemic implementation failure (Table 7). Fidelity rose over the twelve weeks – early sessions (1–4) averaged 15–16, middle sessions (5–8) 16–18 and late sessions (9–12) 17–19 – indicating a facilitator and student learning curve. Four components were hardest to deliver consistently: written individual reflection before discussion (omitted in $\sim 30\%$ of sessions), broad participation of $\geq 80\%$ of students (missed in $\sim 35\%$), recording of behavioural-change commitments (omitted in $\sim 30\%$), and non-evaluative lecturer feedback (present in only $\sim 75\%$). Because fidelity varied across classes (means 16.42–17.78), it is examined descriptively at the class level; formal covariate adjustment and an adequately powered Condition $\times$ Fidelity moderation test lie beyond the twelve-cluster design and are identified as future work (§ 3.6, § 5.7).

Table 7
Mean fidelity score by faculty.

Faculty	Mean score (/22)	Mean %	Category	Range
FPIPS (Social Studies Education)	17.14	77.9%	Moderate	16.08–17.92
FPBS (Language Education)	17.14	77.9%	Moderate	16.33–17.67
FIP (Education Sciences)	16.42	74.6%	Moderate	15.50–16.92
FPMIPA (Mathematics and Science Education)	16.97	77.1%	Moderate	16.17–17.92
FPTI (Information Technology)	17.78	80.8%	High	17.33–18.17
FPOK (Sport and Health Education)	16.66	75.7%	Moderate	16.33–17.08
Overall	17.0	77.3%	Moderate	15.50–18.17
<i>Note. Each faculty contributed one intervention class; values are that class’s mean across its 12 sessions, and Range is the minimum–maximum session score across those 12 sessions. Bands: High $\geq 80\%$; Moderate 60–79%; Low $< 60\%$.</i>				

4.6 Two contrasting pathways of character change

Reflexive thematic analysis produced seven themes, each contrasting a characteristic intervention-group pathway with a comparison-group pathway and each highly consistent across the six faculties (Table 8).

4.6.1 Theme 1 – Learning orientation: transformational experience versus declarative knowledge

Intervention participants recalled learning as active, dialogic and case-based (‘made me think, instead of just listening’, P-006-E), whereas comparison participants recalled declarative knowledge about rules – the ITE Law and hoax penalties (P-004-K) – rather than any verification practice they had carried out.

4.6.2 Theme 2 — Information-verification practice: from social heuristics to lateral reading

Intervention participants described active verification as an internalised habit — opening multiple tabs, comparing sources, checking authors or domains (P-006-E; P-023-E; P-087-E) — while comparison participants relied on social heuristics such as share counts ('I judge it by how many people are sharing', P-015-K); identical situational barriers did not produce an active strategy, so what differs is the formed disposition and skill, not circumstance.

4.6.3 Theme 3 — Sharing behaviour: from impulsive to reflective

The most dramatic shift was from 'just forwarding' to 'thinking first': intervention participants named a conscious, often automatic pause to weigh accuracy and impact, especially on ethnicity, religion and campus politics (P-006-E, P-081-E), whereas comparison participants reported little change and, though some recognised their sharing as impulsive, lacked a mechanism to change it (P-045-K).

4.6.4 Theme 4 — Ethical dilemmas: truth versus social relationships

Facing truth versus another's feelings, intervention participants corrected errors while preserving dignity through private, gently worded messages (P-006-E, P-041-E), whereas comparison participants tended to stay silent to protect social comfort, the decision shaped by relational closeness (P-094-K, P-076-K, P-039-K).

4.6.5 Theme 5 — Response to error: from silent avoidance to open accountability

On realising they had shared falsehood, intervention participants moved toward public clarification, apology and notifying those exposed (P-006-E, P-063-E), whereas comparison participants quietly deleted posts without clarification out of embarrassment or uncertainty (P-004-K, P-082-K).

4.6.6 Theme 6 — Barriers to application: external versus internal

Intervention participants located barriers externally — time pressure, paywalled journals, hard-to-reach sources (P-006-E) — having the will and skill but facing situational friction, whereas comparison participants located them internally, as laziness or low confidence (P-004-K, P-039-K).

4.6.7 Theme 7 — Self-reflection on change: transformation versus stagnation

Intervention participants explicitly saw themselves as changed cognitively, affectively and behaviourally ('I don't get upset as easily when someone believes a hoax', P-058-E), whereas comparison participants reported little change, 'still much the same' recurring (P-004-K, P-039-K, P-057-K).

4.6.8 Lecturer triangulation

Interviews with six lecturers (three teaching both conditions) triangulated the student pattern: those teaching both contrasted an intervention group 'much more used to asking for evidence before accepting a claim' with a comparison group still inclined to accept the majority opinion (D-FPIPS-01), while intervention-only lecturers noted domain-specific shifts (D-FIP-01). Lecturers identified two implementation challenges consistent with § 4.5.1 — preparation load across twelve in-depth sessions (D-FPIPS-01) and disciplinary integration in text-oriented fields (D-FPBS-01) — and proposed discipline-specific curricular pathways.

Table 8
Summary of the seven thematic contrasts.

Theme	Intervention pathway	Comparison pathway	Cross-faculty consistency
Learning orientation	DL – active, dialogic, case-based	KL – passive, lecture-based, rote	Very consistent
Information verification	VA – lateral reading, multi-source checking	VP – social heuristics, virality	Very consistent
Sharing behaviour	JR – reflective pause, impact-oriented	BI – impulsive, driven by content appeal	Very consistent
Ethical dilemmas	KE – empathetic, private correction	HK – conflict avoidance, silence	Very consistent
Response to error	AT – open accountability	PD – silent avoidance	Very consistent
Barriers	HS – external / situational	HM – internal / motivational	Very consistent
Self-reflection	TD – acknowledged transformation	SD – acknowledged stagnation	Very consistent

4.7 Observed behavioural trajectories in the intervention classes

Participant observation of the 18 intervention students (classes EA, EB, EC) at Sessions 1, 6 and 12 corroborated the themes with directly observed change (Table 9). Every code except cognitive dilemma rose from the early to the late session, the largest gains (all + 33.3 pp) in click restraint, moral courage, reflexive correction and collaborative behaviour. At the domain level, epistemic codes moved from emerging to internalised, moral-emotional codes from situational to integrated, and collaborative codes from passive to mature. By the late session 44.4% of intervention participants showed lateral reading and 33.3% click restraint – neither observed in the comparison classes – and 38.9% showed structured collaboration; case-based dilemmas featured in every intervention session but no comparison session, while one-directional instruction characterised 83.3% (5 of 6) of comparison sessions.

Table 9
Distribution of behavioural codes across observation points (intervention, N = 18).

Code	Early session	Mid-session	Late session	Change
LR – Lateral reading	3 (16.7%)	6 (33.3%)	8 (44.4%)	↑ +27.8 pp
CR – Click restraint	0 (0%)	3 (16.7%)	6 (33.3%)	↑ +33.3 pp
TS – Trust shift	0 (0%)	4 (22.2%)	5 (27.8%)	↑ +27.8 pp
EM – Expressed empathy	4 (22.2%)	5 (27.8%)	6 (33.3%)	↑ +11.1 pp
MC – Moral courage	0 (0%)	3 (16.7%)	6 (33.3%)	↑ +33.3 pp
DC – Cognitive dilemma	3 (16.7%)	4 (22.2%)	2 (11.1%)	↓ -5.6 pp*
RC – Reflexive correction	0 (0%)	4 (22.2%)	6 (33.3%)	↑ +33.3 pp
CB – Collaborative behaviour	1 (5.6%)	5 (27.8%)	7 (38.9%)	↑ +33.3 pp
<i>Note. The decline in DC at the late session is interpreted as resolution rather than disengagement: participants appeared to have developed strategies for handling informational confusion, such that the dilemma no longer surfaced as an unresolved cognitive obstacle.</i>				

4.8 Integrating quantitative and qualitative evidence: a joint display

Table 10 integrates the strands, mapping each theoretical pillar (§ 2.7) onto the construct it feeds, the statistical result and the qualitative mechanism, and classifying each row — following Fetters, Curry and Creswell [64] — as convergence, expansion or informative discordance. The interpretation of this pattern is developed in the Discussion (§ 5.1, § 5.4).

Table 10
Integrated joint display: theoretical pillars, quantitative results and qualitative mechanisms.

Theoretical pillar	Quantitative construct & result	Qualitative / observational evidence	Meta-inference
Epistemic Scaffolding	CORA: interaction ns, $F(2,680) = 1.84$, $p = .160$; posttest $g = .12$ ns; convergent analyses $p \approx .07 - .08$ (§ 4.2.2)	LR rose 16.7%→44.4% and CR 0%→33.3% across sessions (Table 9); Theme 2 (VA pathway) documents lateral reading as a spontaneous, internalised habit; portfolio entries record the process directly (§ 4.6.2)	Expansion. The qualitative record shows a genuinely emerging, internalised skill that a short ($\alpha = .67$, $\omega = .79$), single-occasion performance test may be underpowered to register statistically; the qualitative strand raises confidence in the direction of the CORA trend without itself resolving its statistical uncertainty.
Dialogic Deliberation	MSC-D: interaction sig. at posttest, $F(2,680) = 4.29$, $p = .014$, $g = .32^{**}$; not sig. at follow-up, $g = .14$ ns; only digital empathy sig. at dimension level ($p = .034$) (§ 4.2.3)	EM rose 22.2%→33.3% and MC 0%→33.3% (Table 9); Theme 4 (KE, empathetic correction) and Theme 5 (AT, open accountability) are the most vivid intervention-group narratives; interviews were conducted at a single, near-posttest occasion (§ 3.5)	Convergence at posttest, with a durability asymmetry. The dimension driving the quantitative gain (digital empathy) matches the theme most strongly evidenced qualitatively. But the qualitative interviews, timed near posttest, cannot speak to whether participants' sense of transformation itself fades by follow-up in the way the MSC-D score does — a methodological gap for future waves of interviewing.
Reflective Digital Action	DCS: sig., $F(2,680) = 8.33$, $p < .001$, within-group retained, between-group trend at FU ($g = .18$ ns). CPS: sig. and largest, $F(2,680) = 11.50$, $p < .001$, durable to FU ($g = .26^{*}$) (§ 4.2.1, § 4.2.4)	RC rose 0%→33.3% and CB 5.6%→38.9% (Table 9); Theme 5 (open accountability) and Theme 7 (acknowledged transformation) are the most cross-faculty-consistent themes; EA-01's trajectory culminates in collective, forward-looking advocacy leadership (§ 4.7)	Strong convergence, and the pattern explains the DCS/CPS durability gap. CPS-relevant material (collaborative, forward-looking advocacy) is described as consolidated into ongoing identity; DCS-relevant material (situational click-restraint under peer pressure, Theme 6) is explicitly framed by participants as still contingent on situational barriers — a plausible qualitative explanation for why the DCS between-group advantage narrows at follow-up while the CPS advantage does not.

4.9 Boundary conditions: implementation fidelity, deviant cases and equity of effect

The moderator analysis (§ 4.4) showed the effect was demographically equitable; the qualitative strand identifies a complementary boundary condition — how consistently the intervention was delivered, and how individuals responded under comparable delivery. Combining fidelity scores (§ 4.5.1) with class-level behavioural and interview evidence shows a pattern consistent with fidelity moderating change (Table 11): the highest-fidelity class (FPTI, 17.78/22) showed the most mature lateral reading and moral courage, whereas the lowest-fidelity class (FIP, 16.42/22) showed the least mature lateral-reading and collaborative behaviour — motivating a formal, adequately powered Condition × Fidelity test as future work (§ 3.6, § 5.7).

Table 11

Fidelity linked to dominant observed behaviour and interview confirmation, selected classes.

Class	Faculty	Mean fidelity	Dominant observed behaviour	Interview confirmation	Integrated interpretation
FPTI-E	FPTI	17.78 (High)	LR (high), CR (consistent), MC (emergent)	"I've made a habit of checking a source's reputation before believing any offer" (P-081-E)	High fidelity is associated with internalised lateral reading and moral courage.
FPIPS-E	FPIPS	17.14 (Moderate)	TS, EM, CB (mature)	"I posted a clarification and apologised openly" (P-006-E)	Moderate-high fidelity is associated with a trust shift and open accountability.
FPOK-E	FPOK	16.66 (Moderate)	LR (moderate), EM (emergent)	"I look for a BPOM label and related research first" (P-099-E)	Moderate fidelity still produces change, but less consistently than the high-fidelity class.
FPBS-E	FPBS	17.14 (Moderate)	Not in observed subsample; self-reported passive verification	"I'm still in the process of learning" (P-034-E – deviant case)	An individual deviant case (surface learning) despite otherwise moderate class fidelity.

Two deviant cases enrich the account. P-034-E (intervention, FPBS) matched the comparison pattern on five of seven dimensions – a surface-learning pattern underscoring that the intervention is not one-size-fits-all – while P-088-K (comparison, FPTI) showed a genuine but reactive instinct toward open accountability without the lateral-reading skill to prevent the error, illustrating that values without a cognitive method yield good intentions but not resilient competence. Together, § 4.4 and § 4.9 specify two complementary boundary conditions: the effect was equitably distributed across demographic subgroups, but its strength was not uniform across classes or individuals, depending on implementation fidelity and individual factors.

4.10 Summary: cross-source validation of the theoretical model

Read across all sources, the account is internally consistent: the transformation and stagnation pathways held across all seven themes and the six faculties and were independently corroborated by lecturers teaching both conditions; the observed behavioural codes rose from early to late session in the intervention classes while the comparison classes plateaued or declined; and cross-source triangulation – observation, interview, portfolio and instruments – converges on the same three mechanisms (Epistemic Scaffolding, Dialogic Deliberation, Reflective Digital Action), each evidenced by at least three of the four sources, with quantitative confirmation tracking qualitative strength (§ 4.8). The mechanistic interpretation of this convergence is taken up in the Discussion (§ 5.4, § 5.5).

5 Discussion

This study asked whether, and how, a twelve-week deeper-learning intervention could strengthen online civic character, testing a model that treats civic online reasoning, moral self-concept and civic purpose as distinct but pedagogically integrated facets (§ 2.7). The intervention produced significant condition × time effects on three of the four instruments and a favourable trend on the fourth, but did so unevenly, tracing an interpretable gradient (Table 6). We interpret that gradient, then consider what it implies for the integrative model, for equity and fidelity, and for theory, practice and measurement.

5.1 A differential gradient of online civic character

The central result is not that the model simply worked, but that it worked differentially, in a theoretically meaningful order. Motivational and dispositional constructs moved most and lasted longest: civic purpose showed the largest and only durable between-group effect (Hedges $g = 0.36$ at posttest, 0.26 at follow-up), and digital citizenship improved substantially at posttest ($g = 0.35$) with its within-group gain retained. The evaluative self-perception of moral self-concept shifted immediately ($g = 0.32$) but converged by follow-up, and demonstrated civic online reasoning – the sole performance-based measure – yielded only a positive, non-significant trend ($g = 0.12$). Two regularities organise the pattern: dispositional and motivational constructs moved more than demonstrated performance, and forward-looking purpose was sustained better than self-perception. This ordering is the study's central theoretical finding, because it mirrors how the three pillars of the model (§ 2.7) differ in how quickly, and how durably, deeper learning reaches them. The gradient also tracks the observability of each pillar's enactment: Reflective Digital Action, feeding digital citizenship and civic purpose, carried the strongest behavioural-observation evidence and produced the strongest, most durable effects, whereas the most performance-based and least directly observable pillar – Epistemic Scaffolding, feeding civic online reasoning – produced only a non-significant trend, a divergence between statistical and qualitative signal taken up in detail in § 5.3.

5.2 The durable gains: civic purpose and participatory citizenship

Civic purpose was the construct the intervention shifted most strongly and most durably ($F(2,680) = 11.50, p < .001$), and the only one whose between-group advantage survived to the eight-week follow-up while the comparison group's smaller initial gain disappeared entirely (§ 4.2.4). This fits the pillar that fed it – Reflective Digital Action – being enacted through authentic, data-based advocacy rather than discrete exercises: purpose, once folded into a student's sense of who they are online, is comparatively self-sustaining, as the EA-01 trajectory toward collective advocacy leadership illustrates (§ 4.7). The finding extends Kim et al.'s [10] evidence on civic purpose to a Global-South, higher-education setting, echoes accounts linking civic-education quality to participatory motivation [35] and Bennett and colleagues' [1] actualising/dutiful distinction, and is mechanistically consistent with Costa and Oliver's [33] claim that co-producing digital practices builds ownership. The baseline CPS loading non-invariance bears on metric precision, not on the robust direction of an effect corroborated by equivalent pretest means and by both ANCOVA and mixed-model analyses (§ 4.2.4).

Digital citizenship improved significantly and immediately ($g = 0.35$) and retained its within-group gain, but the between-group advantage narrowed to a non-significant trend by follow-up ($g = 0.18$) and was concentrated in the participatory dimensions – internet political activism and local/global awareness – and absent for technical skills, critical perspective and networking (§ 4.2.1). Deeper learning thus moved civic agency rather than procedural skill, consistent with the concern that conventional digital-citizenship scales capture conduct and skill more than character [2, 13]. The follow-up attenuation is explained, not contradicted, by the qualitative record: students framed their residual click-restraint lapses as situational – paywalled sources, peer pressure to respond quickly – rather than motivational (Theme 6, § 4.6.6), so the behaviours most exposed to environmental friction eroded once classroom support was withdrawn, whereas identity-level purpose did not (§ 4.8).

5.3 Immediate, transient and underdetected: moral self-concept and civic online reasoning

Moral self-concept responded to the intervention immediately ($g = 0.32$) but not durably, converging by follow-up, with only digital empathy individually significant at the dimension level (§ 4.2.3). This profile suits the construct: a self-image can be revised quickly under intensive ethical reflection but is sensitive to continuing feedback, so without booster experiences to reconfirm it after the twelve-week dose the initial shift regressed. The contrast with civic purpose is instructive – purpose is a commitment to a direction and comparatively self-sustaining once adopted,

whereas moral self-concept is an evaluative self-perception that needs ongoing reinforcement to consolidate. The behavioural record shows genuine change during the programme (expressed empathy rose 22.2%→33.3% and moral courage 0%→33.3%; Table 9), but the single near-posttest interview wave cannot track whether that sense of transformation fades as the score does (§ 4.8).

Civic online reasoning, the only measure not to reach significance ($g = 0.12$), is the one a quantitative-only reading would most mislead, and three considerations qualify it. First, the analysis was underpowered for so small an effect – the minimum detectable effect was near $d = 0.30$ under the realised near-zero intraclass correlation (§ 3.2) – so a true $g \approx 0.12$ would surface only as a trend, and convergent ANCOVA, mixed-model, gain-score and Mann–Whitney analyses agreed at $p \approx .07-.08$ (§ 4.2.2). Second, the instrument trades sensitivity for feasibility, gaining objective, access-neutral scoring at the cost of a lower ceiling and modest reliability ($\alpha = .67$; § 3.3). Third, and decisively, the observational and documentary evidence shows the skill genuinely emerging (Table 9; Theme 2 and portfolio entries), so the joint display reads this row as expansion (§ 4.8). The null is therefore best understood as a measurement-and-power artefact rather than evidence that reasoning did not improve, converging with wider evidence that well-designed civic-online-reasoning instruction produces measurable gains [7, 17, 36].

5.4 An integrated but empirically separable construct

A key psychometric result frames the whole analysis: the four constructs are only weakly intercorrelated (HTMT .13–.31) and a second-order “online civic character” factor fits poorly (CFI = .35), so the model’s facets are empirically near-orthogonal (§ 4.1). This clarifies rather than undermines the model. Integration is pedagogical and processual – deeper learning is the common bridge that engages epistemic virtues, moral emotions and civic actions within the same authentic tasks – not a claim that the facets collapse into a single latent trait. The differential gradient (§ 5.1) is exactly what distinct developmental targets, reached at different rates, should produce; and the two axial pathways – transformation versus stagnation – recurring across all seven themes and six faculties and corroborated by lecturers who taught both conditions (§ 4.6), describe the cyclical mechanism (cognition → emotion → action → reflection) posited in § 2.7. The practical corollary is that online civic character must be measured and cultivated facet by facet [2, 11, 31].

5.5 Equity, fidelity and individual variation

Two boundary conditions frame the effects. First, the benefit was statistically uniform across sex, socioeconomic status and digital access – of twenty-four moderator tests, none survived false-discovery correction and every interaction explained under 2% of the variance in gains (§ 4.4); in a Global-South setting where digital-citizenship outcomes vary with resources [4, 32, 34], raising lower- and higher-resource students comparably rather than widening existing gaps is a substantive equity result. Second, delivery mattered: fidelity averaged 77.3%, all six intervention classes fell in the moderate-to-high band, and higher-fidelity delivery co-occurred with more consistent lateral reading, the emergence of moral courage and structured collaboration (§ 4.5.1; Table 11). The four hardest-to-deliver components – protected written-reflection time, broad ($\geq 80\%$) participation, systematic recording of commitments, and non-evaluative feedback – are the levers for the next iteration. The deviant cases confirm that the model is not one-size-fits-all: P-034-E engaged without internalising change, while P-088-K showed a values-based ethical instinct without the reasoning method to act on it proactively (§ 4.9), motivating a formal, adequately powered Condition $\times$ Fidelity test.

5.6 Implications for theory, practice and measurement

Theoretically, the gradient adds a durability dimension to the model: deeper learning reaches forward-looking purpose and participatory disposition faster and more durably than objective reasoning performance, which needs more extensive practice to register on a brief test, or evaluative self-concept, which needs sustained reinforcement to

consolidate. Practically, integration should be discipline-specific — the lecturers proposed concrete host courses across the six faculties (§ 4.6.8), and contextual relevance to a student's field reinforced internalisation (§ 4.6.1); durability should be engineered through spaced booster activities (for moral self-concept) and scaffolds that reduce the situational friction — paywalls, peer pressure — students named as their remaining barrier (for digital citizenship); and fidelity should be supported by lightening the preparation load and protecting the hardest-to-deliver components. For measurement, the study advances a validated four-instrument portfolio with established longitudinal invariance, while exposing the reasoning-assessment trade-off: the accessible, key-scored CORA could be paired with a short adaptive or performance-based module to raise its ceiling without restoring the burden that defeated the earlier version.

5.7 Limitations and future directions

Several limitations qualify these findings. The evidence comes from a single Indonesian teacher-education university, so six-faculty representation aids but does not establish transferability across institutions or cultures, and the twelve-cluster design leaves cluster-level inference resting on limited degrees of freedom. Although all four instruments were fully scalar-invariant across occasions, the civic-purpose scale showed baseline loading non-invariance and DCS only partial scalar invariance between groups, so those between-group effects — though supported by equivalent pretest means under randomisation — are interpreted with caution. The CORA result is best read as an underpowered trend rather than a confirmed null, and the moderator analysis tested only three pre-specified demographic factors, leaving unmeasured moderators possible. The qualitative strand was gathered at a single near-posttest wave, limiting how directly it speaks to the follow-up attenuation seen for MSC-D and DCS, and observation covered only three of twelve sessions and three of six classes, so transient episodes between observation points may be missed — partly offset by the denser portfolio record and retrospective interviews. Finally, the deviant cases confirm a non-uniform effect whose interaction between fidelity and individual factors has not yet been tested in a formal, adequately powered way.

Replication with more classes, a follow-up qualitative wave, a more sensitive reasoning measure, and an adequately powered Condition $\times$ Fidelity test are therefore the priorities for extension; these refine rather than qualify the core result, which is robust across the quantitative and qualitative sources.

6 Conclusion

This study set out to test whether online civic character can be cultivated pedagogically, rather than merely regulated, in a digitally saturated Global-South university, using a model that treats civic online reasoning, moral self-concept and civic purpose as distinct but pedagogically integrated facets. A twelve-week deeper-learning intervention delivered to twelve intact classes ($N = 342$) differentially strengthened online civic character: civic purpose showed the largest and only durable gain; digital citizenship improved significantly but with a between-group advantage that narrowed by follow-up; moral self-concept shifted immediately but not durably; and civic online reasoning — the sole performance-based measure — produced a favourable but statistically inconclusive trend that the observational and documentary evidence nonetheless shows to reflect a genuinely emerging skill. Seven cross-faculty themes and directly observed behavioural change explain this gradient through two contrasting pathways — transformation in the intervention group and stagnation in the comparison group — and the benefit was distributed equitably across sex, socioeconomic status and digital access.

Three contributions follow. Empirically, the study provides rare Global-South, higher-education evidence that character-oriented, deeper-learning pedagogy can move the civic dispositions that safety-oriented, procedural digital-citizenship instruction leaves untouched. Theoretically, it shows that an integrative model of online civic character is best understood as pedagogically rather than latently integrated: its facets are empirically separable and develop at

different rates and durabilities, so they must be targeted, and measured, in their own right. Methodologically, it demonstrates the value of a genuinely integrated mixed-methods design, in which a joint display turns an ambiguous quantitative null into an interpretable, qualitatively substantiated trend and locates the mechanisms behind each statistical effect. For practice, the implication is not that online conduct should go ungoverned, but that governance alone is not enough: with discipline-specific integration, engineered durability and well-supported implementation, universities can help students become citizens who reason, empathise and act online — not merely users who comply. Confirming the durability of these effects, and realising them at scale, is the work the next iteration of this design is intended to take up.

Declarations

Funding

This research was funded by a Faculty grant from Universitas Pendidikan Indonesia (Rp 37,500,000).

Competing interests

The authors declare that they have no competing interests.

Ethics approval and consent to participate

All procedures involving human participants were performed in accordance with the ethical standards of the institutional research committee and with the 1964 Declaration of Helsinki and its later amendments. The study was reviewed and approved by the Research Ethics Committee of Universitas Pendidikan Indonesia. Participation was voluntary and without coercion; written informed consent to participate was obtained from all individual participants included in the study.

Consent for publication

All participant data are reported anonymously. Informed consent for the publication of anonymised quotations and data was obtained from the participants.

Availability of data and materials

The datasets generated and analysed during the current study are not publicly available in order to protect participant privacy, but are available from the corresponding author (warlim_isya@upi.edu) on reasonable request.

Authors' contributions

W.W. conceived and designed the study, developed the theoretical model and the instruments, supervised data collection and analysis, and drafted and critically revised the manuscript. M.H. and R.F. contributed to the study design, instrument development and the delivery of the intervention, and to data collection, analysis and interpretation. A.S. and M.I.F. contributed to data collection, coding and analysis, and to the preparation of tables and figures. All authors read, critically revised, and approved the final manuscript.

Acknowledgements

The authors thank the students and lecturers across the six faculties of Universitas Pendidikan Indonesia who took part in this study.

References

1. Bennett WL, Wells C, Freelon D. Communicating civic engagement: contrasting models of citizenship in the youth web sphere. *J Commun.* 2011;61(5):835–56. <https://doi.org/10.1111/j.1460-2466.2011.01588.x>.
2. Choi M, Glassman M, Cristol D. What it means to be a citizen in the internet age: development of a reliable and valid digital citizenship scale. *Comput Educ.* 2017;107:100–12. <https://doi.org/10.1016/j.compedu.2017.01.002>.
3. Chen LL, Mirpuri S, Rao N, Law N. Conceptualization and measurement of digital citizenship across disciplines. *Educ Res Rev.* 2021;33:100379. <https://doi.org/10.1016/j.edurev.2021.100379>.
4. Choi M, Cristol D. Digital citizenship with intersectionality lens: towards participatory democracy driven digital citizenship education. *Theory Pract.* 2021;60(4):361–70. <https://doi.org/10.1080/00405841.2021.1987094>.
5. Pangrazio L, Sefton-Green J. Digital rights, digital citizenship and digital literacy: what's the difference? *J New Approaches Educ Res.* 2021;10(1):15–27. <https://doi.org/10.7821/naer.2021.1.616>.
6. Kahne J, Bowyer B. Educating for democracy in a partisan age: confronting the challenges of motivated reasoning and misinformation. *Am Educ Res J.* 2017;54(1):3–34. <https://doi.org/10.3102/0002831216679817>.
7. McGrew S. Learning to evaluate: an intervention in civic online reasoning. *Comput Educ.* 2020;145:103711. <https://doi.org/10.1016/j.compedu.2019.103711>.
8. Black R, Walsh L, Waite C, Collin P, Third A, Idriss S. In their own words: 41 stories of young people's digital citizenship. *Learn Media Technol.* 2022;47(4):524–36. <https://doi.org/10.1080/17439884.2022.2044848>.
9. Örtengren A. Philosophical underpinnings of digital citizenship through a postdigital lens: implications for teacher educators' professional digital competence. *Educ Inf Technol.* 2024;29:4253–85. <https://doi.org/10.1007/s10639-023-11965-5>.
10. Kim S, Lin T-J, Glassman M, Anderman E. Educating early adolescents for a sustainable future with digital civic learning: moral self-concept as a developmental catalyst linking civic competencies and civic purpose. *Dev Sci.* 2026. <https://doi.org/10.1111/desc.70203>.
11. Warlim W, Effendi R, Fitria R, Ibrahim MM, Hadiati DN. Cross-faculty analysis of AI-enhanced civic character education on digital citizenship development. *Front Educ.* 2026;10:1690799. <https://doi.org/10.3389/educ.2025.1690799>.
12. Ribble M. Digital citizenship in schools. 3rd ed. Portland (OR): International Society for Technology in Education; 2015.
13. Choi M. A concept analysis of digital citizenship for democratic citizenship education in the internet age. *Theory Res Soc Educ.* 2016;44(4):565–607. <https://doi.org/10.1080/00933104.2016.1210549>.
14. Richardson JW, Martin F, Sauers N. Systematic review of 15 years of research on digital citizenship: 2004–2019. *Learn Media Technol.* 2021;46(4):498–514. <https://doi.org/10.1080/17439884.2021.1941098>.
15. Breakstone J, Smith M, Connors P, Ortega T, Kerr D, Saavedra A. Students' civic online reasoning: a national portrait. *Educ Res.* 2021;50(8):505–15. <https://doi.org/10.3102/0013189X211017495>.
16. Wineburg S, Breakstone J, McGrew S, Smith MD, Ortega T. Lateral reading on the open internet: a district-wide field study in high school government classes. *J Educ Psychol.* 2022;114(5):893–909. <https://doi.org/10.1037/edu0000740>.

17. McGrew S, Breakstone J. Civic online reasoning across the curriculum: developing and testing the efficacy of digital literacy lessons. *AERA Open*. 2023;9. <https://doi.org/10.1177/23328584231176451>.
18. Asimakopoulos G, Antonopoulou H, Mitropoulos I, Halkiopoulos C. What makes digital citizenship fragile: a review of the social mechanisms underlying democratic participation. *Societies*. 2026;16(2):70. <https://doi.org/10.3390/soc16020070>.
19. Harrison T. Virtual reality and character education: learning opportunities and risks. *J Moral Educ*. 2024;53(2):219–39. <https://doi.org/10.1080/03057240.2023.2206553>.
20. Peart M, Cubo-Delgado S, Gutiérrez-Esteban P. Exploring the role of digital and socio-civic skills for promoting youth participation and digital citizenship. *Eur J Educ Res*. 2022;11(2):697–708. <https://doi.org/10.12973/eu-jer.11.2.697>.
21. Levy BLM, Busey CL, Cuenca A, Evans RW, Halvorsen A-L, Ho L-C, et al. Social studies education research for sustainable democratic societies: addressing persistent civic challenges. *Theory Res Soc Educ*. 2023;51(1):1–46. <https://doi.org/10.1080/00933104.2022.2148087>.
22. Valdez JP, Li L, Datu JAD. Digital citizenship can prevent online cheating and promote academic delay of gratification. *New Dir Teach Learn*. 2025. <https://doi.org/10.1002/tl.70007>.
23. Vallès-Peris N, Domènech M. Digital citizenship at school: democracy, pragmatism and RRI. *Technol Soc*. 2024;76:102448. <https://doi.org/10.1016/j.techsoc.2023.102448>.
24. Karsgaard C. Democracy and planetary politics: Achille Mbembe and futures of digital citizenship education for life. *Crit Stud Educ*. 2024;65(5):460–75. <https://doi.org/10.1080/17508487.2024.2307573>.
25. Harris A, Johns A. Youth, social cohesion and digital life: from risk and resilience to a global digital citizenship approach. *J Sociol*. 2020;57(2):394–411. <https://doi.org/10.1177/1440783320919173>.
26. McGrew S. Internet or archive? Expertise in searching for digital sources on a contentious historical question. *Cogn Instr*. 2022;40(4):488–516. <https://doi.org/10.1080/07370008.2021.1908288>.
27. Winarno KK, Indrawadi J. Authentic assessment in online learning to develop students' character. *Kasetsart J Soc Sci*. 2023;44(4):1091–100. <https://doi.org/10.34044/j.kjss.2023.44.4.13>.
28. Prasetyo WH, Sumardjoko B, Muhibbin A, Naidu NBM, Muthali'in A. Promoting digital citizenship among student-teachers: the role of project-based learning in improving appropriate online behaviors. *Particip Educ Res*. 2023;10(1):389–407. <https://doi.org/10.17275/per.23.21.10.1>.
29. Polizzi G. Internet users' utopian/dystopian imaginaries of society in the digital age: theorizing critical digital literacy and civic engagement. *New Media Soc*. 2021;25(6). <https://doi.org/10.1177/14614448211018609>.
30. Fajri I, Budimansyah D, Komalasari K. Digital citizenship in civic education learning: a systematic literature review. *J Ilm Peuradeun*. 2022;10(3):833–54. <https://doi.org/10.26811/peuradeun.v10i3.755>.
31. Shi G, Chan KK, Lin XF. A systematic review of digital citizenship empirical studies for practitioners. *Educ Inf Technol*. 2023;28:3953–75. <https://doi.org/10.1007/s10639-022-11383-z>.
32. Lu C, Gu MM. A systematic review and meta-analysis of factors and outcomes of digital citizenship among adolescents. *Asia Pac J Educ*. 2025;45(4):1130–45. <https://doi.org/10.1080/02188791.2023.2296352>.
33. Costa C, Oliver M. Young people and digital literacy learning: co-producing critical citizenship practices. *Pedagogy Cult Soc*. 2025. <https://doi.org/10.1080/14681366.2025.2593611>.
34. Prasetyo WH, Sari BI, Patmisari, Sa'diyah H, Naidu NBM, Prasetyo E. Using Rasch model analysis for assessing psychometric properties of digital citizenship in Indonesian students. *J e-Learn Knowl Soc*. 2022;18(3):96–110. <https://doi.org/10.20368/1971-8829/1135684>.

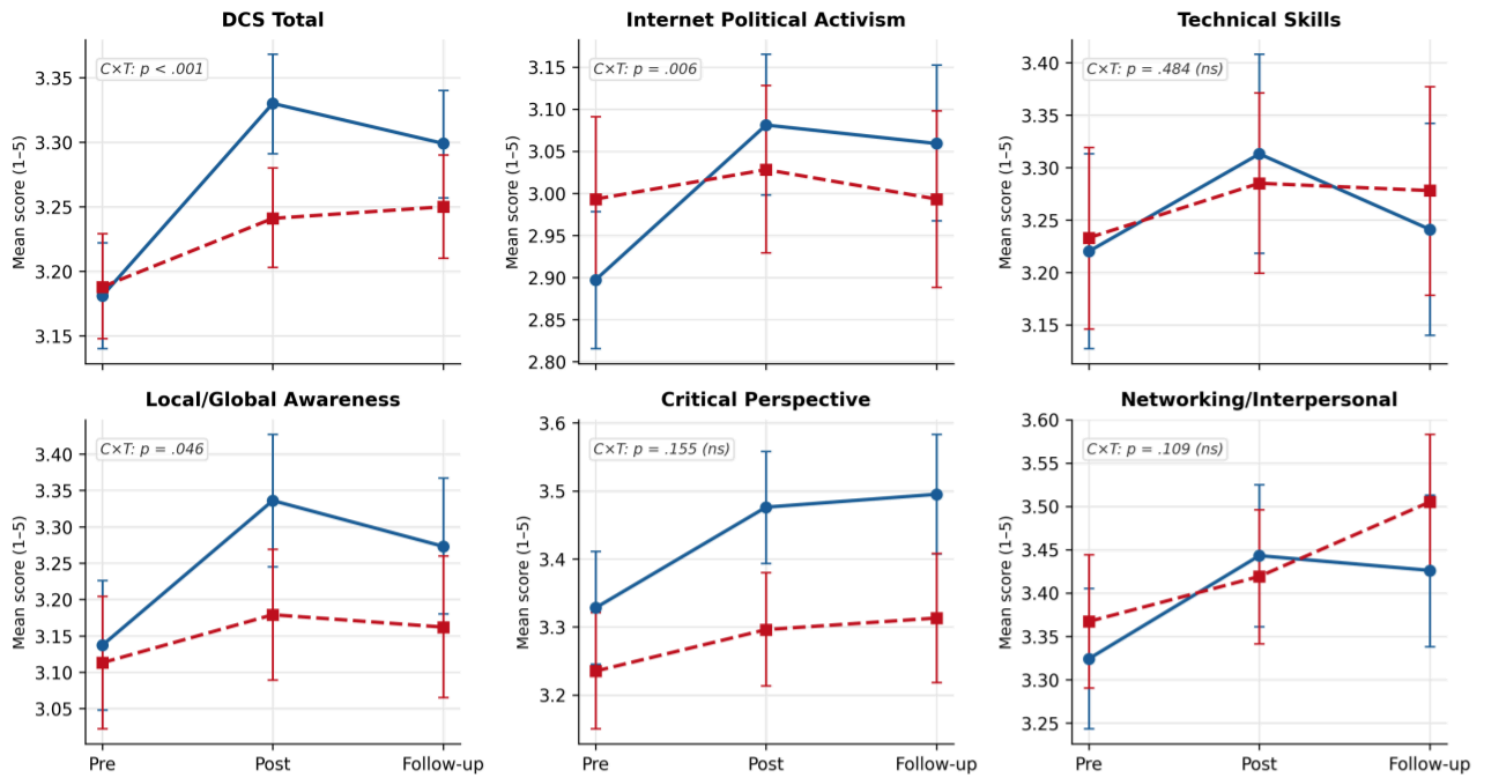
35. Alscher P, Ludewig U, McElvany N. Civic education, teaching quality and students' willingness to participate in political and civic life: political interest and knowledge as mediators. *J Youth Adolesc.* 2022;51:1886–900. <https://doi.org/10.1007/s10964-022-01639-9>.
36. McGrew S, Smith M, Breakstone J, Ortega T, Wineburg S. Improving university students' web savvy: an intervention study. *Br J Educ Psychol.* 2019;89(3):485–500. <https://doi.org/10.1111/bjep.12279>.
37. Westheimer J, Kahne J. What kind of citizen? The politics of educating for democracy. *Am Educ Res J.* 2004;41(2):237–69. <https://doi.org/10.3102/00028312041002237>.
38. Kahne J, Hodgin E, Eidman-Aadahl E. Redesigning civic education for the digital age: participatory politics and the pursuit of democratic engagement. *Theory Res Soc Educ.* 2016;44(1):1–35. <https://doi.org/10.1080/00933104.2015.1132646>.
39. Choi M, Cristol D, Gimbert B. Teachers as digital citizens: the influence of individual backgrounds, internet use and psychological characteristics on teachers' levels of digital citizenship. *Comput Educ.* 2018;121:143–61. <https://doi.org/10.1016/j.compedu.2018.03.005>.
40. Harrison T, Polizzi G. (In)civility and adolescents' moral decision making online: drawing on moral theory to advance digital citizenship education. *Educ Inf Technol.* 2022;27:3277–97. <https://doi.org/10.1007/s10639-021-10710-0>.
41. Cleofas JV, Labayo CC. Youth netizens as global citizens: digital citizenship and global competence among undergraduate students. *Front Commun.* 2024;9:1398001. <https://doi.org/10.3389/fcomm.2024.1398001>.
42. Arkorful V, Salifu I, Arthur F, Nortey SA. Exploring the nexus between digital competencies and digital citizenship of higher education students: a PLS-SEM approach. *Cogent Educ.* 2024;11(1):2326722. <https://doi.org/10.1080/2331186X.2024.2326722>.
43. DiGiacomo DK, Hodgin E, Kahne J, Trapp S. Civic education in a politically polarized era. *Peabody J Educ.* 2021;96(3):261–74. <https://doi.org/10.1080/0161956X.2021.1942705>.
44. BinJwair A, Bingimlas K. Students' practice of digital citizenship values at Prince Sattam Bin Abdulaziz University. *SAGE Open.* 2024;14(2). <https://doi.org/10.1177/21582440241245197>.
45. O'Reilly M, Levine D, Batchelor R, Adams S. Digital ethics of care and digital citizenship in UK primary schools: children as interviewers. *J Child Media.* 2024;18(4):585–604. <https://doi.org/10.1080/17482798.2024.2394932>.
46. Naval C, Villacís JL, Ibarrola-García S. The transversality of civic learning as the basis for development in the university. *Educ Sci.* 2022;12(4):240. <https://doi.org/10.3390/educsci12040240>.
47. Park H, Fulton K, Anderman EM. Cultivating civic competencies through immersive inquiry: a digital-age approach to fourth graders' disciplinary thinking and argumentation. *J Soc Stud Res.* 2024;49(2). <https://doi.org/10.1177/23522798241301436>.
48. Lozano-Díaz A, Fernández-Prados JS. Educating digital citizens: an opportunity to critical and activist perspective of Sustainable Development Goals. *Sustainability.* 2020;12(18):7260. <https://doi.org/10.3390/su12187260>.
49. Pangrazio L, Bunn A. Assessing the privacy of digital products in Australian schools: protecting the digital rights of children and young people. *Comput Educ Open.* 2024;6:100187. <https://doi.org/10.1016/j.caeo.2024.100187>.
50. Guath M, Nygren T. Civic online reasoning among adults: an empirical evaluation of a prescriptive theory and its correlates. *Front Educ.* 2022;7:721731. <https://doi.org/10.3389/educ.2022.721731>.
51. Pellegrino JW, Hilton ML, editors. Education for life and work: developing transferable knowledge and skills in the 21st century. Washington (DC): National Academies; 2012. <https://doi.org/10.17226/13398>.
52. Assante GM, Popa NL, Momanu M. How personal values and critical dispositions support digital citizenship development in higher education students. *Front Psychol.* 2022;13:990518. <https://doi.org/10.3389/fpsyg.2022.990518>.

53. Kahne J, Bowyer B. Can media literacy education increase digital engagement in politics? *Learn Media Technol.* 2019;44(2):211–24. <https://doi.org/10.1080/17439884.2019.1601108>.
54. Mirra N, McGrew S, Tynes B, et al. Expanding digital citizenship education to address tough issues. *Phi Delta Kappan.* 2022;103(5). <https://doi.org/10.1177/00317217221079976>.
55. Mulyono B, Affandi I, Suryadi K, Darmawan C. Online civic engagement through social media: an analysis of Twitter big data. *Cakrawala Pendidik.* 2023;42(1). <https://doi.org/10.21831/cp.v42i1.54201>.
56. Hemer KM, Kappus A. Assessment of civic learning and democratic engagement. *New Dir High Educ.* 2022. <https://doi.org/10.1002/he.20417>.
57. Kim M, Choi D. Development of Youth Digital Citizenship Scale and implication for educational setting. *Educ Technol Soc.* 2018;21(1):155–71. <https://www.jstor.org/stable/26273877>.
58. Tadlaoui-Brahmi A, Çuko K, Alvarez L. Digital citizenship in primary education: a systematic literature review describing how it is implemented. *Soc Sci Humanit Open.* 2022;6(1):100348. <https://doi.org/10.1016/j.ssaho.2022.100348>.
59. Creswell JW, Plano Clark VL. *Designing and conducting mixed methods research.* 3rd ed. Thousand Oaks (CA): SAGE; 2018.
60. Cohen J. *Statistical power analysis for the behavioral sciences.* 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988.
61. Lincoln YS, Guba EG. *Naturalistic inquiry.* Beverly Hills (CA): SAGE; 1985.
62. Field A. *Discovering statistics using IBM SPSS statistics.* 5th ed. London: SAGE; 2018.
63. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol.* 2006;3(2):77–101. <https://doi.org/10.1191/1478088706qp063oa>.
64. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs — principles and practices. *Health Serv Res.* 2013;48(6 Pt 2):2134–56. <https://doi.org/10.1111/1475-6773.12117>.
65. Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct Equ Model.* 2007;14(3):464–504. <https://doi.org/10.1080/10705510701301834>.
66. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res.* 1981;18(1):39–50. <https://doi.org/10.1177/002224378101800104>.
67. Malhotra NK, Dash S. *Marketing research: an applied orientation.* 7th ed. New Delhi: Pearson; 2016.

Figures

Digital Citizenship Scale trajectories by condition (pretest, posttest, 8-week follow-up)

— Experimental (n = 172) — Comparison (n = 170)

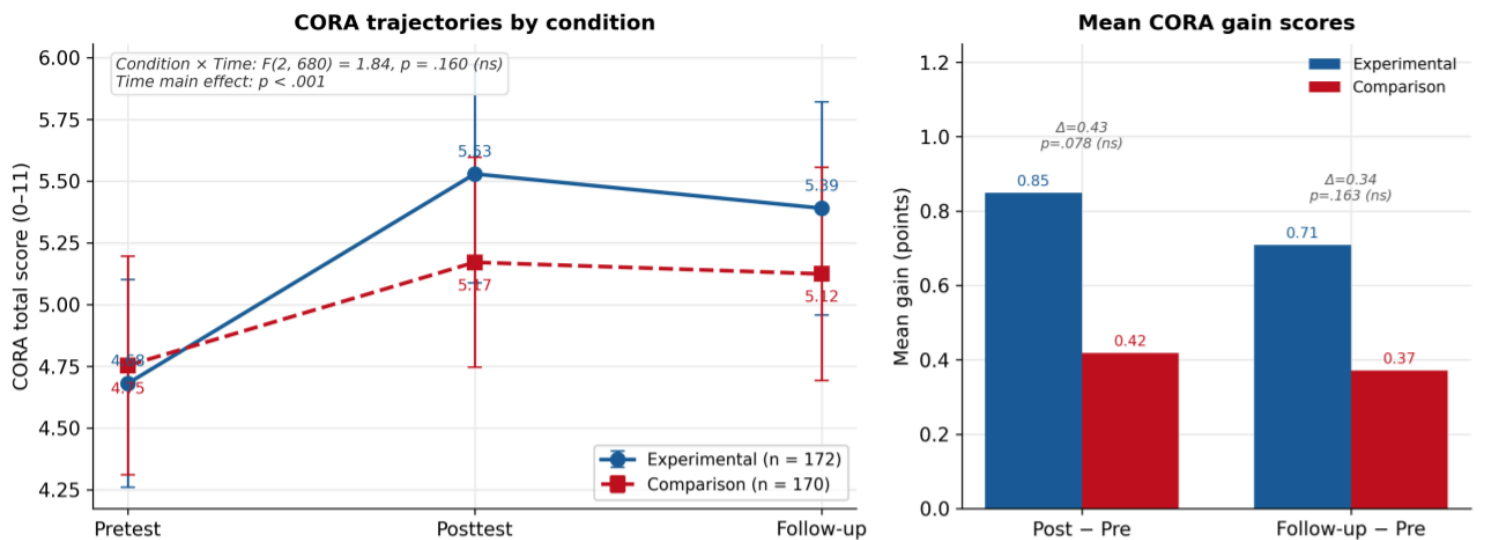


Error bars = 95% CI. CxT = Condition x Time interaction (2x3 mixed ANOVA, Greenhouse-Geisser corrected).

Figure 1

Digital citizenship (DCS) trajectories by condition across pretest, posttest and follow-up.

Civic Online Reasoning Assessment (CORA): a positive but non-significant intervention trend



Error bars = 95% CI. CORA total = sum of 8 tasks (0-11). No significant Condition x Time interaction on any analytic framework.

Figure 2

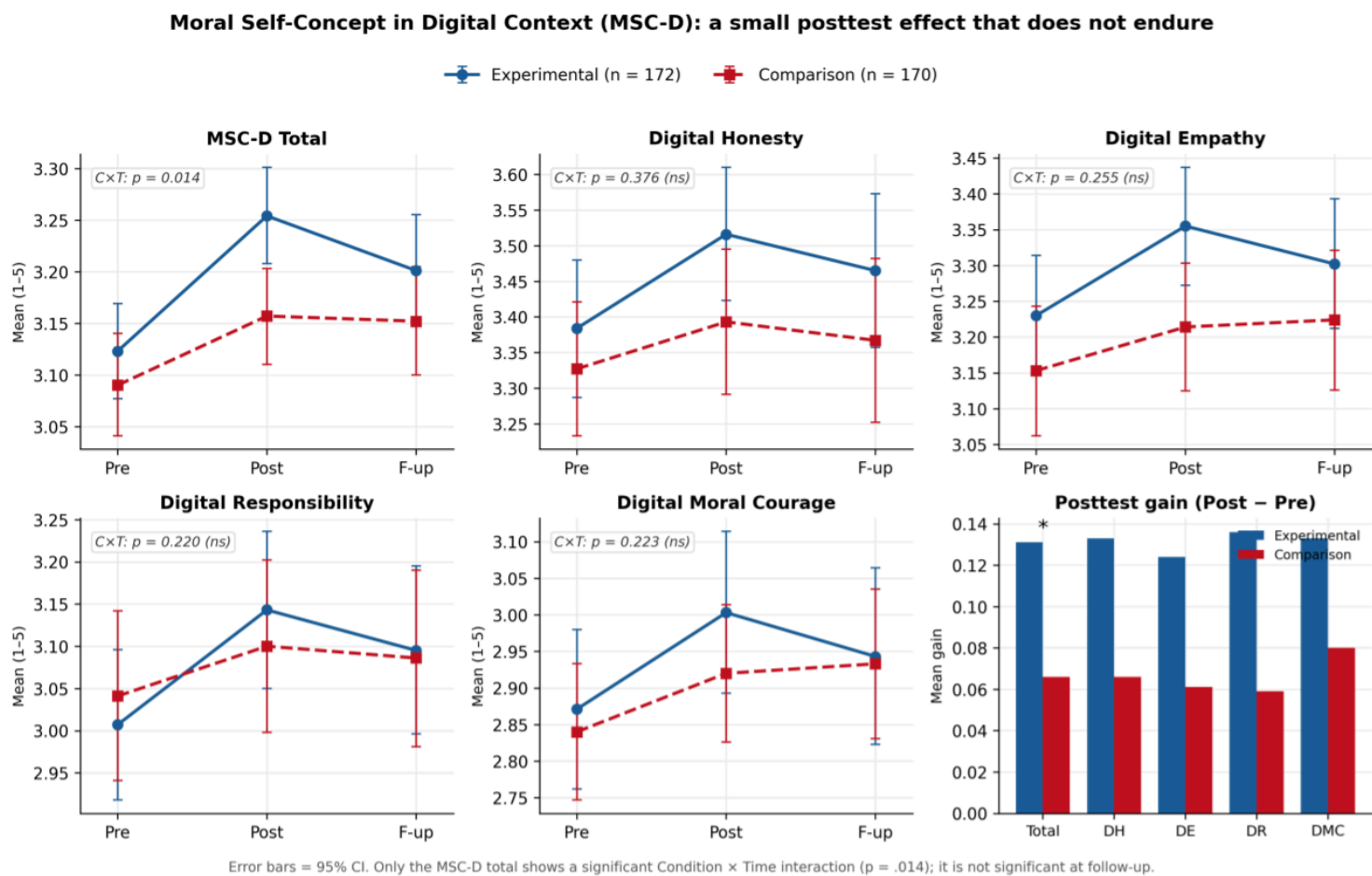


Figure 3

Moral self-concept (MSC-D) trajectories by condition across pretest, posttest and follow-up.

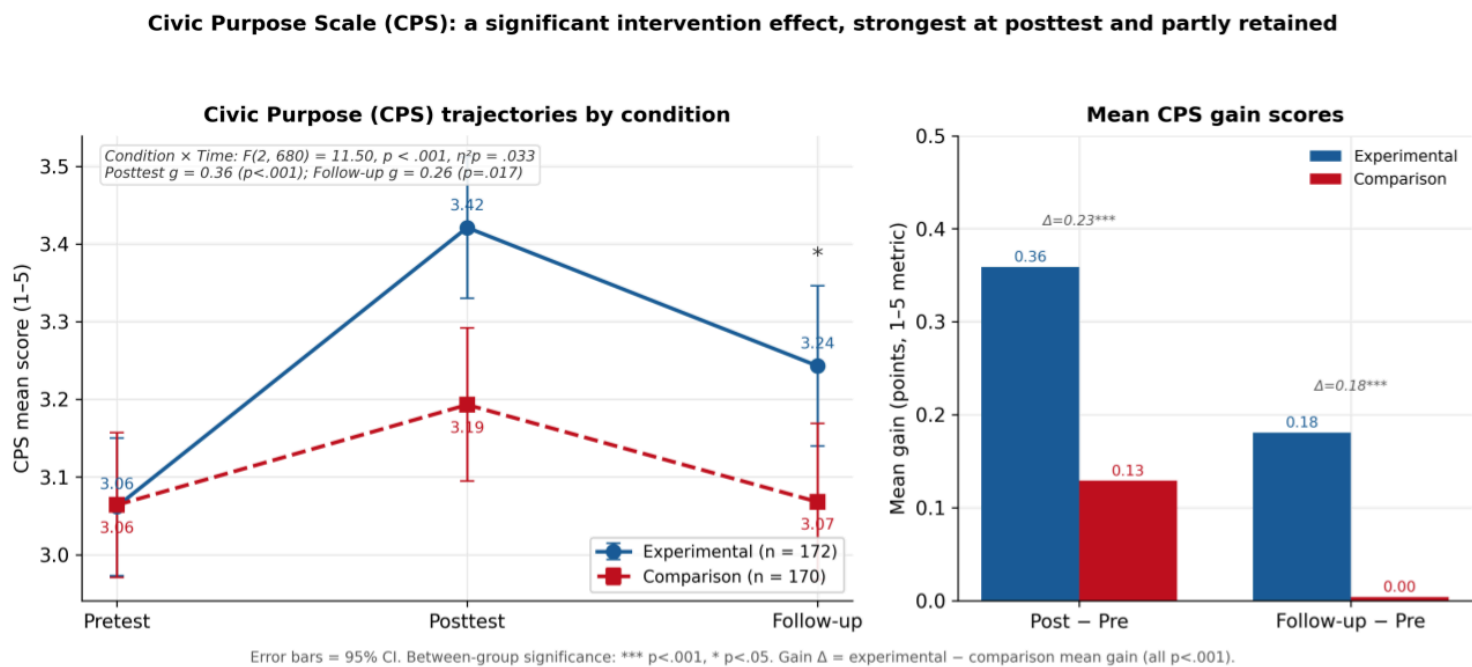


Figure 4

Civic purpose (CPS) trajectories by condition across pretest, posttest and follow-up.

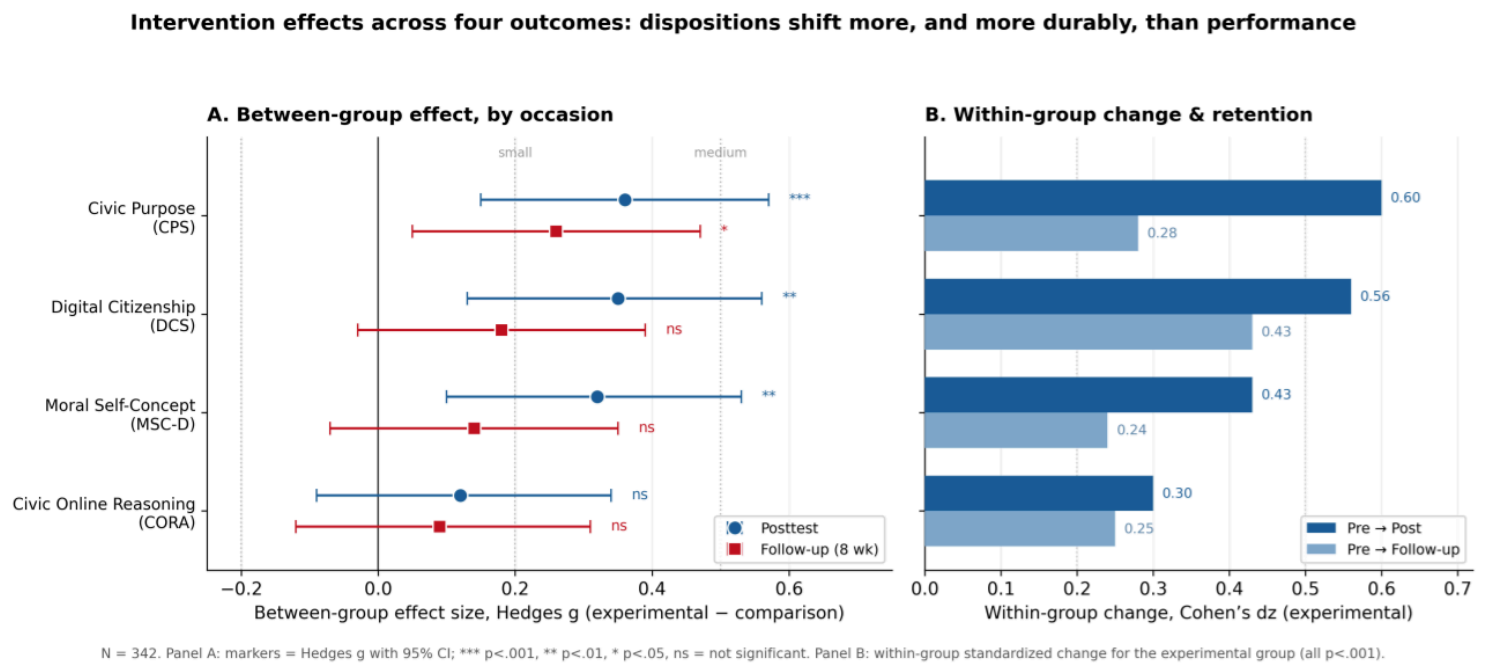
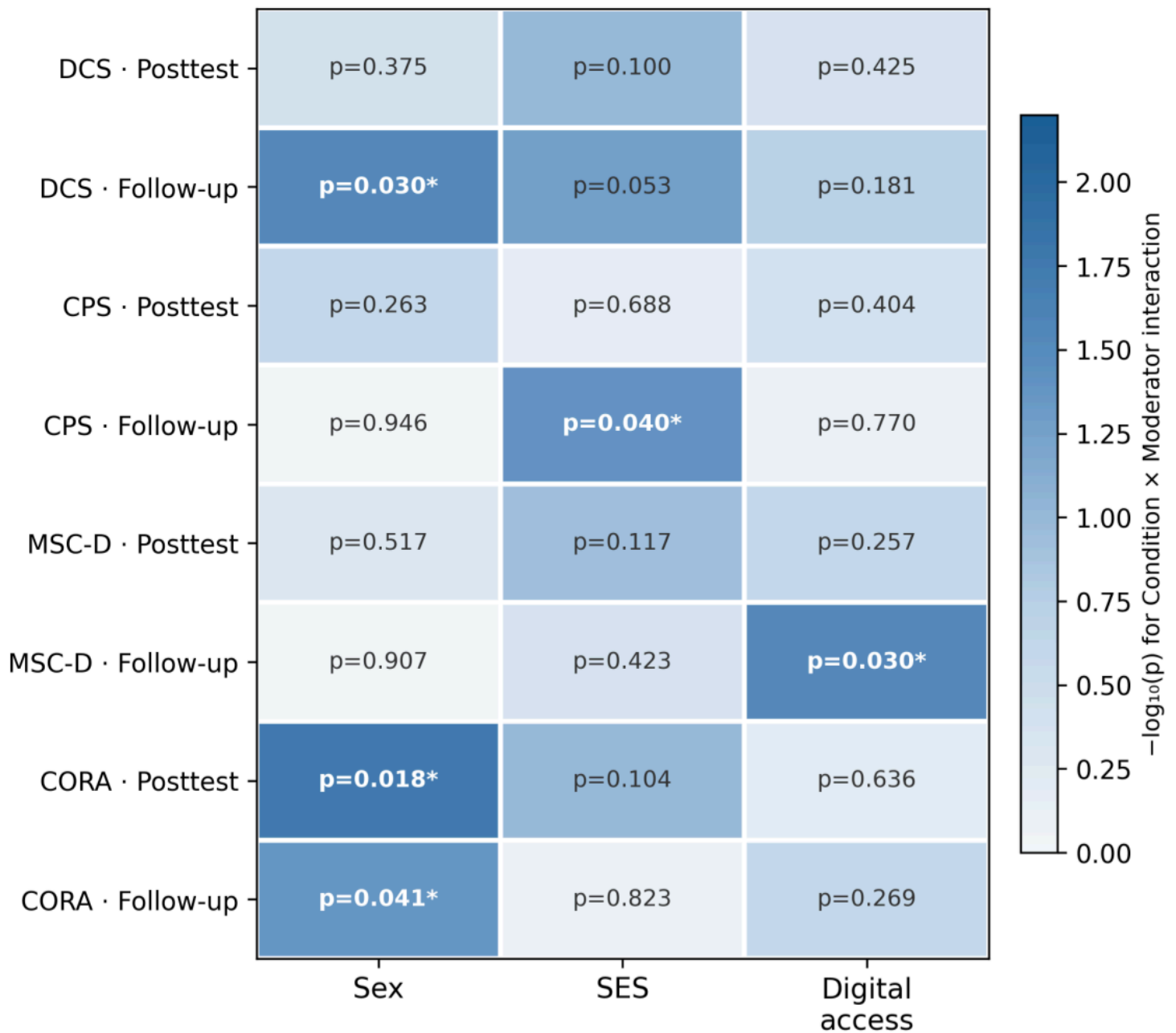


Figure 5

Cross-instrument synthesis. (A) Between-group effect sizes (Hedges g, 95% CI) at posttest and follow-up; (B) within-group standardised change for the intervention group. Dispositions and purpose shift more, and more durably, than demonstrated performance.

Moderation of the intervention effect **(Condition × Moderator interaction on gain scores)**



* raw $p < .05$ (5 of 24 tests). None survive FDR correction; all interaction $\eta^2 p^2 < .02$.
 Effects were statistically uniform across sex, SES and digital access.

Figure 6

Condition × moderator interaction p-values across instruments and occasions. Darker = smaller p; * marks raw $p < .05$ (5 of 24 tests). None survive FDR correction; all interaction $\eta^2 p < .02$.