Journal of Accounting, Ethics & Public Policy, JAEPP
e-ISSN 2956-8390; ISSN1089652X Vol. 27 No. 1 (2026)

From Awareness to Action: Operationalizing Accounting Ethics

Sam Islam, PhD
Department of Accounting,
Martha & Spencer Love School of Business, Elon University, Elon, NC

Doi: 10.60154/jaepp.2026.v27n1p60

Submitted: 05 January 2026
Accepted: 20 May 2026
Published: 22 May 2026

Cite As:
Islam, S. (2026). *From Awareness to Action: Operationalizing Accounting Ethics*. Journal of Accounting, Ethics & Public Policy, JAEPP, 27 (1), 60.
https://doi.org/10.60154/jaepp.2026.v27n1p60

Abstract

Accounting programs operate under a recurring tension: AACSB outcomes and IFAC IES 4 require evidence of ethics learning, but curricular integration often lacks the instructional mechanisms to produce such evidence. This study examines a four-component instructional sequence, E⁴P (Exposure, Engagement, Enactment, and Partnership), and the assurance-of-learning artifacts it generates. The paper combines an integrative scoping review with a course-embedded pilot in Accounting Information Systems (n=121) to test a mechanism-aligned measurement suite comprising ethical sensitivity vignettes, DIT-style reasoning, the Hurtt skepticism scale, and a Giving Voice to Values (GVV) performance rubric. The instruments are organized through a Measurement and Scoring Framework that specifies constructs, scoring rubrics, mastery thresholds, and archivable artifacts suitable for assurance-of-learning review. Within the pilot setting, skepticism shifts modestly over a single term, while ethical sensitivity and GVV-rehearsed enactment show larger short-run gains. The paper interprets these patterns as within-student movement consistent with the E⁴P sequence, not as evidence of durable trait change or institution-wide effectiveness, and identifies the conditions under which such artifacts can support accreditation review across courses.

Keywords: Accounting ethics education; assurance of learning; Giving Voice to Values; professional skepticism; moral reasoning; scoping review

I. Introduction

Ethics education has been central to accounting's public-interest mandate since the modern professional structure took shape, and a sequence of accounting scandals from Enron and WorldCom through more recent restatement episodes has kept the question of whether universities can prepare graduates to recognize and act on ethical concerns active in both research and policy debate (McPhail, 2001; Mintz, 2021). That question now sits within explicit expectations from accrediting and professional bodies: AACSB emphasizes learning outcomes and societal impact, IFAC's IES 4 specifies learning outcomes for professional values, ethics, public interest, and skepticism, and the IESBA Code supplies a principles-based anchor for classroom cases and assessment (AACSB, 2020, 2025; IFAC, 2021; IESBA, 2024). At the same time, U.S. ethics licensure requirements vary across jurisdictions, complicating curricular design and documentation (Horne et al., 2022).

Integration practices remain uneven. Program surveys report growth in standalone ethics courses alongside variability in embedded coverage and assessment (Robinson et al., 2020). Historical syntheses reach a similar conclusion: integration without explicit outcomes, rubrics, and assurance tends to be shallow, while standalone courses provide depth but risk isolation from technical contexts (Nguyen and Dellaportas, 2020; Mintz, 2021). Prior empirical work shows that structured analysis and reflective discussion can raise ethical sensitivity and moral reasoning in the short run, though connections to implementation and durability are less consistently demonstrated (Dellaportas, 2006; Thomas, 2012). Professional skepticism appears to shift more slowly and seems to require repeated calibration in audit-relevant settings (Hurtt, 2010).

This study takes up that gap by examining a standards-aligned instructional sequence, E⁴P (Exposure, Engagement, Enactment, Partnership), and the assurance-of-learning measurement spine that accompanies it. E⁴P is not advanced as a new moral theory. It translates established accounts of moral functioning and issue-contingent judgment into a set of instructional moves that target a chronic weakness in ethics education: the transition from recognizing and reasoning about dilemmas to rehearsing and communicating values-consistent action under realistic constraints (Rest and Narvaez, 1994; Jones, 1991; Gentile, 2010).

The paper proceeds in two steps. An integrative scoping review first synthesizes accounting-education evidence on ethics pedagogy, mechanisms, and outcomes, and documents the search and screening process (SRP, Appendix A1; EMAP, Appendix A2). A course-embedded AIS pilot then evaluates feasibility and short-run responsiveness using four outcomes aligned to the E⁴P sequence: ethical sensitivity (0–1), DIT-style moral

reasoning, professional skepticism (Hurtt, 2010), and a performance-based Giving Voice to Values rubric (1–4) (MSF, Appendix A3). Results are interpreted as within-student movement consistent with the mechanism, not as claims of durable trait change or institution-wide effectiveness.

The study contributes to the accounting ethics education literature in two ways. First, it specifies an instructional sequence (E⁴P) that maps each component to AACSB outcomes, IES 4 competencies, and the IESBA Code, thereby clarifying the mechanism by which standards-level expectations are delivered at the course level (Robinson et al., 2020; Nguyen and Dellaportas, 2020). Second, it pairs the sequence with a Measurement and Scoring Framework that defines constructs, instruments, scoring rules, and mastery thresholds, producing the archivable artifacts that programs use to document ethics learning for accreditation review (AACSB, 2020, 2025; IFAC, 2021).

II. Conceptual Framework: E⁴P and Study Aims

E⁴P organizes ethics instruction into four components: Exposure to principles and standards, Engagement through structured analysis, Enactment via rehearsed action under realistic constraints, and Partnership with practice. In this framing, awareness and reasoning are plausible mediators between curriculum and outcomes, while perceived moral intensity and ethical climate are expected to moderate effect sizes. The study (i) synthesizes evidence on practices that map to E⁴P, (ii) proposes a standards-aligned implementation pathway that preserves instructor autonomy, and (iii) reports a course-embedded pilot that uses a compact, archivable assessment suite. The sequence implies testable expectations: Exposure and Engagement strengthen recognition and reasoning in the short run, while Enactment more directly influences action-relevant performance within constrained scenarios. Partnership is expected to matter most for durability and workplace transfer, which the present study treats as a question for replication rather than one settled by a single pilot (Dellaportas, 2006; Thomas, 2012; Gentile, 2010).

III. Literature Synthesis and Context

Integration models and common barriers

Programs tend to choose among three designs: integration across technical courses, a standalone course, or a blend. Surveys and historical reviews indicate that integration without explicit outcomes and rubrics can be thin, while standalone courses provide depth but risk isolation unless linked to program-level goals (Robinson et al., 2020; Nguyen and Dellaportas, 2020). Blended “sandwich” structures, which combine early exposure, threaded reinforcement, and capstone synthesis, address both depth and continuity (Armstrong, 1993; Fleming et al., 2009; Hurtt and

Thomas, 2008). Constraints persist: crowded curricula, uneven faculty preparation, and varying state requirements (Loeb, 1988; Horne et al., 2022). Together, these features support a portable design with clear outcomes and lightweight artifacts that can be archived for assurance.

Early work documented wide variation in ethics coverage and time allocation and debated the place of a required course, establishing the baseline from which current integration efforts proceed (Ponemon, 1993; Ahadiat and Mackie, 1993; Madison and Schmidt, 2006; Blanthorne et al., 2007; Armstrong et al., 2003; Klimek and Wenell, 2011). Newer technologies can also widen ethical blind spots, suggesting the need for ongoing calibration as tools and data practices change (Sherif and Mohsin, 2021). Evidence from non-U.S. contexts echoes these themes and points to local constraints and opportunities in integrating ethics (Ahmad, 2015).

What tends to work, and why

Empirical studies link structured analysis and reflective discussion to gains in ethical sensitivity and moral reasoning (Dellaportas, 2006; Thomas, 2012; Mintz, 2019, 2021). Enactment tasks, in particular GVV scripting and rehearsal under realistic constraints, are associated with improved implementation confidence and intent to act (Gentile, 2010; Manly and Ritsema, 2018; Loeb, 2015). Studies examining behavioral intent in early-career contexts suggest that instruction can shape reporting behavior and that mentors and peers influence ethical intentions, complementing classroom design (Mayhew and Murphy, 2009; McManus and Subramaniam, 2009). A virtue-ethics lens points to dispositions that interact with casework and rehearsal (Thorne, 1998). Professional skepticism generally shifts more slowly and benefits from repeated, audit-relevant calibration (Hurtt, 2010). Active and reflective pedagogies (cases, debates, role-play, journals) reinforce these patterns (Bebeau, 1994; Jagger et al., 2016). The accounting ethics education literature points to a mechanism pathway consistent with E^4P: Exposure and Engagement strengthen recognition and reasoning, Enactment develops action, and Partnership supports transfer.

Comparators and global standards

Law and medicine illustrate structures that make ethics visible and cumulative: clear outcomes, designated assessment points, and audit trails of learning (NCBE, 2025; LCME, 2024). In law, accreditation guidance and professional responsibility requirements are formalized in ABA standards alongside the MPRE (American Bar Association, 2025; NCBE, 2025). In medicine, ethics is reinforced through accreditation and clinical learning, where students describe ethical reflection embedded in routine rounds (LCME, 2024). Engineering has long emphasized codified norms in

undergraduate training (ASEE, 1999). Cross-disciplinary work in accounting draws on these comparators to motivate structured integration and assessment (Liu et al., 2012). For accounting, standards already provide a suitable frame: AACSB outcomes and societal impact, IFAC's IES 4 competencies, and the IESBA Code. International initiatives, including ACCA's Ethics and Professional Skills Module, point to a trend toward practice-proximate simulation and transparent competency claims (ACCA, 2025; O'Leary and Mohamad, 2006). These comparators are not offered as templates for imitation; they support the portability and accountability logic embedded in E⁴P.

Assurance of learning in accounting ethics

Accreditation-driven assurance of learning (AoL) is the practical mechanism through which accounting programs translate ethics expectations into reviewable evidence, and the accounting-ethics AoL literature converges on three design features that distinguish defensible programs from documentation that satisfies process without informing instruction. The first is alignment: AoL evidence is most useful when course-level artifacts map directly to a small number of program-level learning goals that are themselves anchored to AACSB outcomes and IFAC IES 4 competencies, an alignment that the surveyed accounting programs in Christensen, Judd, and Nichols (2011) were still in the process of building. The second is direct measurement: scoring rubrics applied to course-embedded artifacts (cases, written analyses, performance tasks) carry more weight in accreditation review than indirect indicators such as exit surveys, and the recursive process model in Zocco (2011) makes direct rubric-based assessment the operational core of the AoL cycle. The third is closing the loop: evidence is only useful when it feeds curricular adjustment, and the two-cohort case studies in Lawrence, Reed, and Locander (2011) and in Chiu and Fischer (2018) document how rubric scores from one cohort revealed instructional gaps that, once addressed through a redesigned ethics module, produced measurable changes in the next cohort. In parallel, Martinov-Bennie and Mladenovic (2015) show that an integrated ethics component embedded in a first-year accounting unit raised ethical sensitivity, while exposure to a decision framework on its own affected judgment but not sensitivity, an asymmetry that implies AoL designs in accounting ethics need separate measurement of recognition, reasoning, and enactment rather than a single composite ethics score. Read together, these contributions support an AoL design for accounting ethics that pairs explicit goal-instrument mapping with rubric-scored, course-embedded artifacts and treats the assessment results as an instructional signal rather than an accreditation artifact (Rexeisen and Al-

Khatib, 2009). The Measurement and Scoring Framework developed below operationalizes that design for the E⁴P sequence.

IV. Scoping Literature Review Method

Protocol. The synthesis follows the Scoping Review Protocol detailed in Appendix A1 (SRP) and is reported in line with PRISMA-ScR guidance for scoping reviews (Tricco et al., 2018). The protocol is designed to map concepts, instruments, and evidence relevant to the E⁴P sequence rather than to estimate pooled effects.

Databases, time frame, and search strings. Searches covered January 2000 through July 2025 in Scopus, Web of Science, ERIC, and Google Scholar. The search string paired *accounting* with the topical set (*ethics* OR *professional responsibility* OR *integrity*) and the instructional set (*education* OR *curriculum* OR *pedagogy* OR *assessment*), targeting the title, abstract, and keyword fields.

Eligibility criteria. A study was included if it was peer-reviewed and English-language, focused on accounting students or early-career accountants, and reported a learning or behavioral outcome relevant to the E⁴P sequence (ethical awareness or sensitivity, moral reasoning, ethical intent or behavior, or professional skepticism). Conceptual and standards documents from AACSB, IFAC (IES 4), and the IESBA Code were retained to situate program-level expectations (AACSB, 2025; IFAC, 2021; IESBA, 2024). Studies were excluded if the population was non-accounting without an explicit transfer rationale, if they were opinion or commentary without an outcome measure, or if the outcome was not extractable from the reported design.

Screening and data extraction. Records identified across databases were de-duplicated and screened against the eligibility criteria at the title-and-abstract stage, then re-assessed at full text. The full PRISMA-style screening flow—records identified, duplicates removed, titles and abstracts screened, full texts assessed, full texts excluded with reasons, and studies retained for qualitative synthesis—is reported in Table A of Appendix A1. Eligible studies were coded into the Evidence Matrix (EMAP, Appendix A2), which records, for each source, the study setting and design, participants, instruments (e.g., DIT-style reasoning, ethical-sensitivity vignettes, GVV performance, Hurtt skepticism), main effects, durability evidence, and noted moderators (moral intensity, ethical climate), together with each study's mapping onto the E⁴P components. To limit selection bias, the included set was cross-checked against existing syntheses on accounting ethics education, principally Nguyen and Dellaportas (2020).

Acknowledged constraints. The synthesis is integrative rather than meta-analytic and does not estimate pooled effect sizes; the English-

language and peer-reviewed restrictions raise the possibility of language and publication bias; and category coding involves judgment rather than algorithmic adjudication. These constraints are revisited in the Limitations section and motivate the paper's emphasis on a transparent measurement spine in the course-embedded pilot rather than on comparative-effectiveness claims derived solely from the review.

V. Implementation Strategies and Assurance Toolkit

The implementation strategies developed below make the E⁴P sequence operational. Exposure is delivered through brief primers and case pre-reads; Engagement through structured analysis and short simulations; Enactment through GVV rehearsal under realistic constraints; and Partnership through employer-facing artifacts. Each component maps to AACSB outcomes, IES 4 competencies, and relevant IESBA principles. Prior accounting ethics education work points to greater traction when ethics is treated as a curriculum-wide competency rather than a single-course outcome (Nguyen and Dellaportas, 2020; Robinson et al., 2020).

A blended "sandwich" arrangement, in which early exposure to foundational concepts is followed by recurring applications in technical courses and synthesized in a culminating experience, gives students both depth and continuity. Programs that adopt this design can map coverage and depth to AACSB outcomes, IFAC IES 4 learning objectives, and the IESBA Code, clarifying expectations and building an assurance trail without prescribing a single format (Nguyen and Dellaportas, 2020). The toolkit that follows specifies a minimal set of artifacts that allow instructors to preserve autonomy while enabling programs to document coverage, performance, and year-over-year improvement.

Implementation Toolkit and Assurance of Learning

Operationalizing the foregoing requires a lightweight toolkit that instructors can adapt to course needs while maintaining consistency with IES 4 and program-level AACSB outcomes. A workable kit comprises (i) a small set of cases mapped to specific outcomes, (ii) a one-page ethical-sensitivity vignette calibrated for moral-intensity cues with a transparent scoring guide, (iii) a GVV performance rubric that captures awareness, reasoning, voice/courage, and outcome, and (iv) brief prompts that support transfer to internship or project settings. Instructors select the elements that fit their syllabus, and the program aggregates the artifacts to document coverage and learning. The approach is consistent with earlier toolkit-based proposals in accounting and with IFAC's IAESB Ethics Education Toolkit (Dellaportas et al., 2008; IAESB, 2015). More broadly, experiential methods developed in

business ethics, including structured simulations and coached role-play, align with the GVV emphasis on rehearsal (Thorne LeClair and Ferrell, 2000).

Assurance of learning can proceed in a staged way. Programs collect several artifacts each term (for example, one sensitivity score, one reasoning artifact, and one GVV performance score per student) and review summary metrics annually. Faculty then use the evidence to calibrate where ethics prompts sit in technical courses, which cases merit replacement, and whether additional rehearsal is warranted. Partner engagement, such as an internship mentor's two-item rating of professional conduct, adds a modest practice-facing indicator without overburdening employers. The aim is not a rigid template but a traceable link between intended outcomes, instruction, and evidence of student learning. Together, these elements form a light yet auditable pathway in which instructors retain discretion over materials while programs aggregate a small set of artifacts for assurance review.

VI. AIS Ethics Mini-Pilot (Course-Embedded Evaluation)

Sample and Method

A course-embedded pilot evaluates the feasibility and short-run responsiveness of the integration strategies described above. The pilot was conducted across four undergraduate Accounting Information Systems (AIS) sections over three semesters at a single U.S. institution, taught by one instructor, yielding a pooled sample of 121 paired pre/post observations. The instructional sequence combined an early ethics module with threaded prompts in AIS topics (data governance, internal controls, systems change management), culminating in a graded Giving Voice to Values (GVV) enactment in which student teams rehearsed and delivered values-consistent responses to realistic constraints (client pressure, supervisory directives).

Four outcomes were assessed pre/post: ethical sensitivity (O1), operationalized as scenario recognition on a vignette calibrated for moral intensity (0–1 scale); moral reasoning (O2), captured with a DIT-style summary score; professional skepticism (O3), measured with the Hurtt (2010) instrument (1–6); and implementation confidence (O4), a performance score from a four-dimension GVV rubric (awareness, reasoning, voice/courage, outcome; 1–4). The GVV task followed case analysis and required students to script and enact specific actions rather than merely identify issues (Gentile, 2010). Analyses used paired t-tests on difference scores with 95% confidence intervals for the mean difference and Cohen's *d* for paired samples. A priori, mastery thresholds were set at ≥.80 for ethical sensitivity (O1) and ≥3.25 for GVV performance (O4), consistent with rubric-based mastery levels used in related accounting ethics education work. The evaluation was embedded in assurance-of-learning processes aligned with AACSB outcomes and IFAC IES 4 learning objectives.

Table 1: Pre- and Post-instruction Outcomes of AIS Ethics Mini-pilot Study

Outcome	*O1: Ethical Sensitivity*	*O2: Moral Reasoning (DIT-2 proxy)*	*O3: Professional Skepticism (HPS)*	*O4: Implementation Confidence (GVV)*
N (paired)	121	121	121	121
Pre Mean	0.608	25.316	4.774	2.619
Pre SD	0.15	5.079	0.507	0.634
Post Mean	0.764	29.474	4.958	3.165
Post SD	0.157	6.797	0.551	0.574
Mean Δ(Post-Pre)	0.156	4.158	0.185	0.546
95% CI Δ	[0.131, 0.181]	[2.949, 5.367]	[0.12, 0.249]	[0.439, 0.653]
t (df=120)	12.84	6.98	5.83	10.38
p	< .001	< .001	< .001	< .001
Cohen's d (paired)	2.17	1.18	0.98	1.75
% Improved	97.1	88.6	82.9	97.1
Units	Proportion (0-1)	Score	1-6	1-4

This table presents the Pre- and post-instruction outcomes for the AIS ethics mini-pilot study (paired n = 121). Columns report means and standard deviations at each time point, mean change (post – pre), 95% confidence interval for the mean change, paired t with df = 120, p value, Cohen's d (paired), percent of students improving, and scale units.

Results from Pilot Study

Table 1 reports descriptive statistics, mean changes, confidence intervals, t values, and effect sizes. Because two outcomes are measured on bounded scales (ethical sensitivity 0–1, GVV rubric 1–4), standardized mean-change statistics can appear large when pre-intervention dispersion is modest, and instruction produces a relatively uniform shift. For this reason, interpretation emphasizes (i) the raw mean changes and their confidence intervals, (ii) the proportion of students improving, and (iii) mastery rates tied to a priori thresholds, rather than treating Cohen's d as a standalone claim about generalizable magnitude. This framing is consistent with the paper's purpose, which is to evaluate the feasibility of mechanism-aligned measurement and to provide an assurance-of-learning template that can be replicated in other course settings.

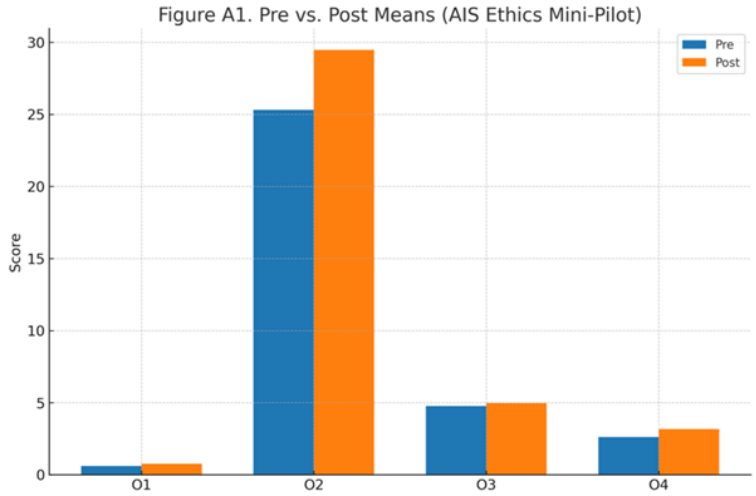

Figure 1: Pre/post means of AIS Ethics Mini-pilot Study for O1–O4

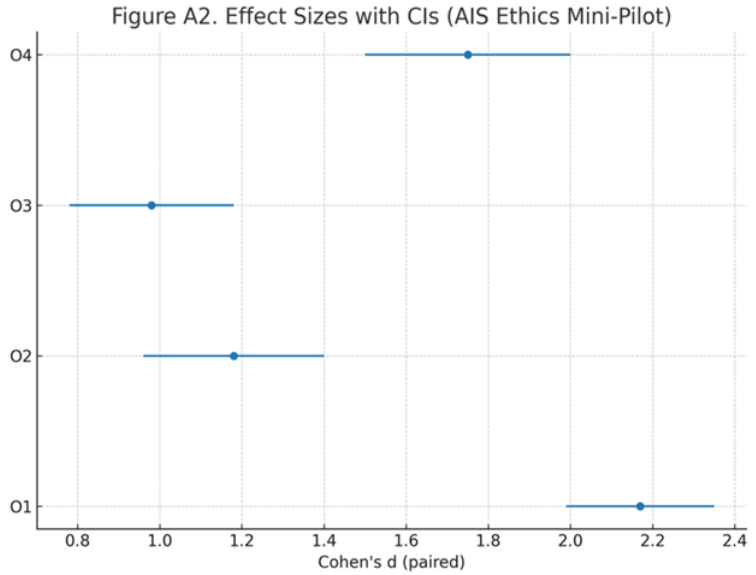

Figure 2: Effect sizes (Cohen's d) of AIS Ethics Mini-pilot Study for O1–O4 with confidence intervals

Ethical sensitivity increased by 0.156 points on the 0–1 scale, 95% CI [0.131, 0.181], $t(120) = 12.84$, $p < .001$, indicating a short-run improvement in students' ability to recognize the presence and contours of an ethical issue in AIS settings. Post-instruction, 48.6% of students met the mastery criterion (≥ .80; Table 3). Moral reasoning improved by 4.16 points on the DIT-style score, 95% CI [2.95, 5.37], $t(120) = 6.98$, $p < .001$, consistent with prior evidence that structured discussion and targeted instruction elevate

principled reasoning in the short run (Dellaportas, 2006; Thomas, 2012). Because DIT-family scores capture moral reasoning schemas rather than ethical behavior, this movement is interpreted as a necessary but insufficient cognitive shift for ethical action. Accordingly, the design pairs reasoning with a performance-based enactment measure (GVV) to better approximate movement from issue recognition to action within the course setting (Elm and Weber, 1994; Marnburg, 2001).[1]

Professional skepticism also increased, albeit more modestly, by 0.185 points on the Hurtt 1–6 scale (95% CI [0.12, 0.25]), $t(120) = 5.83$, $p < .001$. Because the Hurtt measure is commonly interpreted as dispositional, this shift is interpreted as increased activation of skeptical judgment habits in an AIS-relevant learning context rather than as evidence of stable trait change, which would require repeated calibration across multiple courses and settings (Hurtt, 2010; Robinson et al., 2018; Nolder and Kadous, 2018). The largest change is observed on the GVV performance measure: implementation confidence rose by 0.546 on the 1–4 rubric (95% CI [0.439, 0.653]), $t(120)=10.38$, $p<.001$, with 48.6% of students meeting the enactment mastery threshold (≥ 3.25; Table 1; Table 3). Figure 1 displays pre/post means, and Figure 2 summarizes effect sizes. The pattern shows larger short-run gains for enactment (O4) and recognition (O1) than for skepticism (O3), with reasoning (O2) between the two.[2]

Table 2: GVV Sub-dimensions of AIS Ethics Mini-pilot Study

Outcome	*GVV Awareness*	*GVV Reasoning*	*GVV Voice/Courage*	*GVV Outcome*
N (paired)	121	121	121	121
Pre Mean	2.901	2.784	2.55	2.576
Pre SD	0.566	0.593	0.64	0.685
Post Mean	3.296	3.226	3.186	3.086
Post SD	0.675	0.57	0.596	0.793
Mean Δ(Post-Pre)	0.396	0.441	0.637	0.51
95% CI Δ	[0.275, 0.516]	[0.33, 0.553]	[0.521, 0.752]	[0.393, 0.627]
t (df=120)	6.67	8.05	11.19	8.84
p	< .001	< .001	< .001	< .001
Cohen's d (paired)	1.13	1.36	1.89	1.49
% Improved	88.6	88.6	97.1	91.4
Units	1-4	1-4	1-4	1-4

[1] While DIT measures have documented validity evidence, they remain indirect proxies (Thoma & Dong, 2014; Marnburg, 2001).

[2] As a robustness check for distributional assumptions in pre/post comparisons, nonparametric Wilcoxon signed-rank tests were also estimated for each outcome. The direction of effects is unchanged, and inference remains consistent with the paired t-tests [Wilcoxon Z and p-values to be inserted]. The triangulation reduces reliance on a single parametric specification and supports interpreting the observed changes as short-run within-student movement consistent with the E⁴P sequence.

This table presents the GVV performance rubric sub-dimensions (awareness, reasoning, voice/courage, outcome) of the pilot.

Rubric sub-dimension analyses in Table 2 and Figure 3 show improvements across all four GVV dimensions. Mean increases were 0.396 (awareness), 0.441 (reasoning), 0.637 (voice/courage), and 0.510 (outcome), all p <.001. The largest change in voice/courage suggests that students became more adept at articulating concerns and proposing workable, values-consistent alternatives when facing interpersonal and organizational pressure, a capability emphasized by GVV (Gentile, 2010). Gains on the outcome dimension are consistent with greater facility in choosing subsequent actions and anticipating second-order effects in systems-oriented AIS scenarios. Together, these results are consistent with the E⁴P logic: Exposure to standards and dilemmas, Engagement through case analysis, Enactment via rehearsed scripts under realistic constraints, and Partnership with practice through assurance-of-learning artifacts that can be reviewed with employers.

Table 3: Mastery Thresholds of AIS Ethics Mini-pilot Study

	O1 ≥ 0.80 Ethical Sensitivity	O4 ≥ 3.25 GVV Enactment
% of students	48.59	48.55

This table presents the results of the mini-pilot test's mastery rates (O1 ≥ .80 and O4 ≥ 3.25) for students.

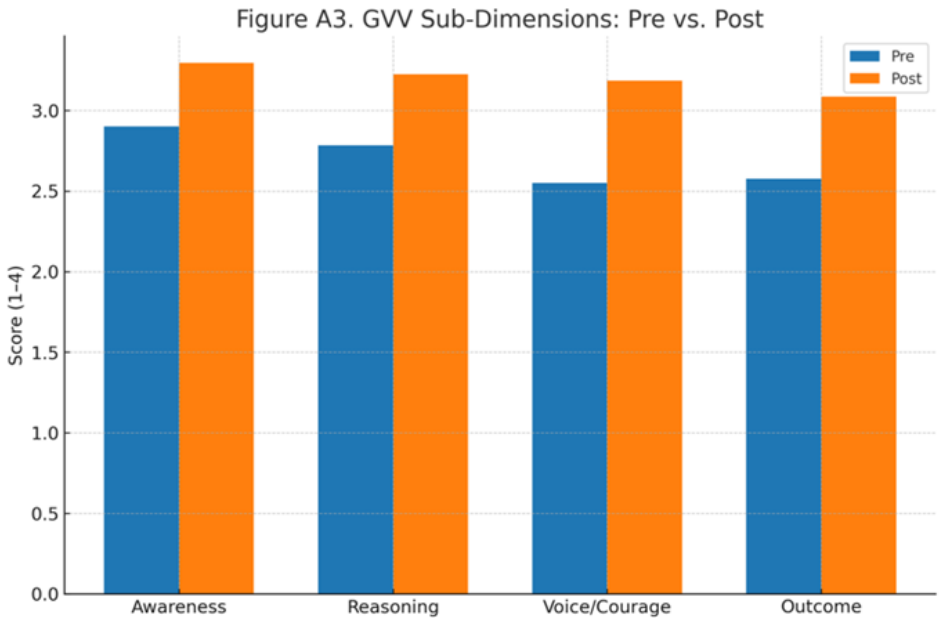

Figure 3: Pre/post means of AIS Ethics Mini-pilot Study for GVV sub-dimensions

Read against the Evidence Matrix (EMAP, Appendix A2), the AIS pattern is consistent with prior accounting education findings: structured analysis and discussion elevate reasoning and recognition (Dellaportas, 2006; Thomas, 2012), while enactment tasks produce larger performance shifts under constraint (Gentile, 2010). Modest movement on the Hurtt skepticism scale is expected over a single term and supports repeated calibration within audit-relevant contexts (Hurtt, 2010). Because the Measurement and Scoring Framework (MSF, Appendix A3) aligns each outcome with a specific instrument and rubric, the assurance trail is feasible, and the standards are consistent.

VII. Discussion: Implications and Bridge to Practice

The pilot pattern, with larger gains in enactment and ethical sensitivity, meaningful gains in reasoning, and smaller but positive movement in professional skepticism, is consistent with the E⁴P sequence and with prior accounting education evidence (Dellaportas, 2006; Thomas, 2012; Hurtt, 2010; Gentile, 2010). Read alongside the literature synthesis and the standards frame, the results suggest a pathway by which classroom activity may translate into enacted, values-consistent responses within course-relevant scenarios. Workplace transfer and durability remain questions for replication across cohorts and institutions. The paper does not claim universal effect sizes from a single context. Rather, the E⁴P sequence clarifies the instructional logic, and the accompanying artifacts provide a documentation trail that programs can reuse for accreditation review. The framing supports a cumulative research agenda in which future studies test scope conditions (cohort differences, institution type, and course placement) while preserving a comparable measurement and documentation spine.

Instructional implications

The findings point to the value of rehearsal. Students who script and practice values-consistent responses under realistic constraints appear better positioned to move from analysis to action, which is the outcome emphasized by the GVV approach (Gentile, 2010). Case analysis and discussion remain important: they strengthen recognition and reasoning, and they provide the context in which enactment can be rehearsed (Dellaportas, 2006; Thomas, 2012). The modest short-run movement in skepticism is consistent with its trait-like features and suggests that audit-relevant calibration is best distributed across courses rather than concentrated in a single module (Hurtt, 2010). In program terms, a blended “early exposure, threaded reinforcement, capstone synthesis” architecture fits these dynamics without prescribing a single course format.

Curricular alignment with standards

AACSB emphasizes outcomes and evidence; IFAC's IES 4 specifies competencies in ethics and skepticism; and the IESBA Code provides a principles-based anchor (AACSB, 2020, 2025; IFAC, 2021; IESBA, 2024). The E⁴P sequence can be situated within this accountability logic, and the E⁴P-Standards Implementation Map (E⁴P-SIM, Appendix A4) makes the mapping explicit: Exposure links to principles and threats, Engagement to case analysis and simulations, Enactment to GVV performance, and Partnership to practice-facing artifacts. The mapping preserves instructor autonomy while making program-level intent and evidence visible. Appendix A5 documents scope limits and durability considerations.

Assessment and assurance of learning

The Measurement and Scoring Framework (MSF, Appendix A3) provides a compact set of indicators (ethical sensitivity, DIT-style reasoning, Hurtt skepticism, and a four-dimensional GVV rubric) that programs can archive and review annually. Mastery thresholds help faculty interpret change and identify where reinforcement is needed. The instruments are proxies for behavior, but their combination captures distinct facets of competence and supports continuous improvement consistent with accreditation expectations (AACSB, 2020, 2025; IFAC, 2021). Programs can use MSF trends to calibrate case difficulty, placement of ethics prompts in technical courses, and the frequency of enactment tasks.

Bridge to practice

A brief internship mentor check-in (two ratings and an open comment) provides an external lens on enactment without imposing undue burden on employers. Employer observations can inform case realism and the timing of ethics prompts, strengthening the Partnership component and supporting early-career socialization around integrity, objectivity, and independence (IESBA, 2024). Collaborative routines with employers are consistent with prior evidence on university–industry partnerships that facilitate bidirectional learning and skill signaling (Ankrah and Al-Tabbaa, 2015). Given state-by-state variation in pre-licensure ethics requirements, portable artifacts and a shared vocabulary for outcomes can ease documentation and mobility (Horne et al., 2022).

Public policy and licensure relevance

State-by-state variation in ethics education and ethics examination requirements creates a documentation problem with policy consequences. Candidate mobility across jurisdictions can depend on whether ethics coverage is recognizable to regulators, boards, and employers. The E⁴P

approach contributes to this discussion by making ethics instruction legible in terms of evidence. Programs can document (i) where ethical principles and threats are introduced (Exposure), (ii) where case-based judgment is evaluated (Engagement), (iii) where action planning and communication are assessed using a performance rubric (Enactment), and (iv) whether workplace supervisors observe early-career enactment behaviors (Partnership). The framing complements debates about how much ethics should be required by shifting attention to how ethics learning can be documented with minimal burden while preserving institutional autonomy. The accountability logic is consistent with a public-interest framing of the profession and offers a bridge between curriculum design and the evidentiary expectations that sit behind accreditation and licensure discussions.

Implementation considerations

Adoption can proceed incrementally. Instructors begin by adding one GVV enactment to an existing case, introducing a short ethical-sensitivity vignette, and capturing a single MSF indicator for assurance-of-learning review. Over time, departments broaden coverage across courses, estimate inter-rater reliability for the GVV rubric, and engage practice partners in periodic review of anonymized artifacts. The staged approach preserves manageable workloads while building a coherent evidence trail.

VIII. Limitations and Directions for Future Research

Several limitations qualify the interpretations advanced in this paper. First, the integrative scoping review emphasizes breadth and transparency rather than formal causal inference. The Scoping Review Protocol (SRP; Appendix A1) documents a PRISMA-style search and screening, but the synthesis remains English-language only and does not estimate pooled effects. Heterogeneity in designs, measures, and contexts precluded meta-analysis and limits claims about comparative effectiveness. Publication bias is also possible, as studies with null results may be underrepresented. These constraints reinforce the paper's aim to align mechanism, pedagogy, and measurement rather than to adjudicate all alternative approaches.

Second, the design is single-institution and single-instructor. While the paired sample (n = 121) is adequate for pre/post inference at the section level, generalization to other institutional contexts requires replication. The pilot pooled matched pre/post observations across four undergraduate AIS sections taught by one instructor over three semesters at one institution. Without a comparison section and a longer follow-up, internal validity and durability cannot be established. Instructor effects, cohort composition, and local curricular sequencing may have contributed to the observed gains. As summarized in the Evidence Matrix (EMAP, Appendix A2), short-run

improvements in recognition and reasoning are common in the literature, while movement in professional skepticism is typically slower and sensitive to audit-relevant context (Dellaportas, 2006; Thomas, 2012; Hurtt, 2010).

Third, the measurement strategy balances feasibility and construct coverage but relies on proxies. The Measurement and Scoring Framework (MSF, Appendix A3) combines an ethical-sensitivity vignette (0–1), a DIT-style reasoning score, the Hurtt skepticism scale (1–6), and a four-dimension GVV performance rubric (1–4). These instruments are widely used, but they remain indirect indicators of real-world behavior. Inter-rater reliability for the GVV rubric was not estimated in this study, and the DIT-style summary was not benchmarked against an external criterion. Mastery thresholds (O1 ≥ .80; O4 ≥ 3.25) were set a priori to improve interpretability, but alternative cut points would yield different mastery rates. These features recommend caution in generalizing magnitudes, even where the pattern of effects is theoretically coherent (Gentile, 2010; Hurtt, 2010). The DIT-style reasoning proxy should be interpreted as a moral reasoning structure rather than ethical behavior, and the literature includes debate about how well moral development scores track enacted ethics in professional contexts (Bay, 2002). When used as part of a triangulated assessment suite rather than as a standalone measure, the DIT remains a defensible indicator of cognitive moral development. Likewise, because professional skepticism is often treated as dispositional, short-run movement is interpreted cautiously as context-linked activation rather than durable trait change (Elm and Weber, 1994; Marnburg, 2001; Robinson et al., 2018; Hurtt, 2010).

Fourth, generalizability beyond AIS warrants caution. The E⁴P sequence (Exposure, Engagement, Enactment, Partnership) was operationalized with AIS-relevant cases and constraints, including systems change and data governance. Transfer to audit, tax, or financial reporting likely depends on domain-specific dilemmas, perceived moral intensity, and the ethical climate into which students transition (Rest and Narvaez, 1994; Jones, 1991). Programs with different accreditation emphases or employer partnerships may implement the E⁴P-Standards Implementation Map (E⁴P-SIM, Appendix A4) in ways that shift emphasis across components.

These limits define a research agenda. Multi-site, course-embedded studies that vary the presence and timing of GVV enactments and incorporate comparison sections can test the incremental contribution of Enactment over structured analysis alone. Longitudinal designs that link classroom artifacts to internship mentor ratings and early-career observations would address durability and transfer, extending the Partnership component. Measurement work that estimates inter-rater reliability of the GVV rubric, calibrates DIT-style summaries to common benchmarks, and examines invariance across course types would strengthen interpretability. Moderator

analyses of ethical climate, incentive structures, and moral-intensity cues can clarify when and for whom E⁴P produces the largest gains (Dellaportas, 2006; Thomas, 2012; Hurtt, 2010).

Conclusion

The paper addresses a recurring challenge in accounting ethics education: programs are expected to prepare students for ethically consequential practice, yet the field often relies on broad aspirations or indirect proxies that are difficult to defend as assurance-of-learning evidence. The integrative scoping review and the course-embedded pilot together translate that ambition into an assessable pathway that separates four complementary outcomes (ethical sensitivity, ethical reasoning, professional skepticism, and values-based enactment) and shows how each can be taught and evaluated within standard accounting curricula. The central contribution is practical. The Measurement and Scoring Framework provides scoring rules, mastery thresholds, and classroom-ready artifacts that programs can use in accreditation documentation and continuous-improvement cycles without rebuilding assessment infrastructure.

Within a single semester, students improve on outcomes that are plausibly responsive to instruction, particularly enactment-related performance, when ethics is treated as a practiced capability rather than a topic to cover. The paper does not overclaim what one-course interventions can accomplish. Trait-like dispositions and workplace behavior are shaped by repeated practice, reinforcement across courses, and organizational context. The framework does not by itself "produce" ethical professionals. It strengthens the conditions under which ethics learning develops and can be documented in defensible ways.

For accounting programs and accreditors, the framework provides a bridge between professional expectations and educational accountability. It supports curriculum mapping to professional standards, documents learning gains in an auditable form, and enables comparability over time as programs iterate on instructional design. For instructors, the design logic is straightforward: teach for action, measure what is responsive to instruction in the short run, and build cumulative reinforcement across the curriculum.

Several limitations point to a focused research agenda. Multi-institution replications are needed to test generalizability across student populations and delivery modes, and longitudinal designs are needed to examine whether enactment gains persist and translate into internship and early-career decision contexts. Future work can also refine construct coverage by triangulating rubric-based performance with behavioral observations, peer assessments, and practitioner evaluations. These steps would deepen the evidence base while preserving the paper's aim of making

accounting ethics education assessable and defensible to the stakeholders who rely on the profession's public-interest commitment.

Conflict of Interest: The author reported no conflict of interest.

Data Availability: All data are included in the content of the paper.

Funding Statement: The author did not obtain any funding for this research.

References:

1. AACSB. (2020). *2020 business accreditation standards (updated June 2023)*. AACSB International.
2. AACSB. (2025). *2020 guiding principles and standards for business accreditation (updated through February 28, 2025)*. AACSB International.
3. ACCA. (2025). *Ethics and Professional Skills Module (EPSM)*. Association of Chartered Certified Accountants.
4. Ahadiat, N., & Mackie, J. J. (1993). Ethics education in accounting: An investigation of the importance of ethics as a factor in the recruiting decisions of public accounting firms. *Journal of Accounting Education*, *11*(2), 243–257. https://doi.org/10.1016/0748-5751(93)90005-4
5. Ahmad, N. L. (2015). Integrating ethics into accounting curriculum: Overview from Malaysian accounting educators. *Geografia: Malaysian Journal of Society and Space*, *11*(6), 87–97.
6. American Bar Association. (2025). *ABA standards for approval of law schools*. https://www.americanbar.org/groups/legal_education/resources/standards/
7. Ankrah, S., & Al-Tabbaa, O. (2015). Universities–industry collaboration: A systematic review. *Scandinavian Journal of Management*, *31*(3), 387–408.
8. Armstrong, M. B. (1993). Ethics and professionalism in accounting education: A sample course. *Journal of Accounting Education*, *11*(1), 77–92.
9. Armstrong, M. B., Ketz, J. E., & Owsen, D. (2003). Ethics education in accounting: Moving toward ethical motivation and ethical behavior. *Journal of Accounting Education*, *21*(1), 1–16.
10. ASEE (American Society for Engineering Education). (1999). *Ethics in Engineering Education*. ASEE.

11. Bay, D. (2002). A critical evaluation of the use of the DIT in accounting ethics research. *Critical Perspectives on Accounting*, *13*(2), 159-177.
12. Bebeau, M. J. (1994). Influencing the moral dimensions of dental practice. In J. R. Rest & D. Narváez (Eds.), *Moral Development in the Professions: Psychology and Applied Ethics* (pp. 133–158). Hillsdale, NJ: Lawrence Erlbaum.
13. Blanthorne, C., Kovar, S. E., & Fisher, D. G. (2007). Accounting educators' opinions about ethics in the curriculum. *Issues in Accounting Education*, *22*(3), 355–390.
14. Chiu, H. H., & Fischer, D. (2018). Implementing assurance of learning: Developing an ethics assessment program for a school of business. *Journal of Business Ethics Education*, *15*, 33–48.
15. Christensen, A. L., Judd, A. J., & Nichols, N. B. (2011). Implementation of assurance of learning plans: An accounting program and individual course analysis. *Journal of Education for Business*, *86*(2), 84–91. https://doi.org/10.1080/08832323.2010.480990
16. Dellaportas, S. (2006). Making a difference with a discrete course on accounting ethics. *Journal of Business Ethics*, *65*(4), 391–404.
17. Dellaportas, S., Jackling, B., & Leung, P. (2008). Ethics education for accounting students—A toolkit approach. *Accounting Education*, *17*(sup1), S27–S43.
18. Elm, D. R., & Weber, J. (1994). Measuring moral judgment: The Moral Judgment Interview or the Defining Issues Test? *Journal of Business Ethics*, *13*(5), 341–355. https://doi.org/10.1007/BF00871762
19. Fleming, D. M., Romanus, R. N., & Lightner, S. M. (2009). The effect of professional context on accounting students' moral reasoning. *Issues in Accounting Education*, *24*(1), 13–30.
20. Gentile, M. C. (2010). *Giving Voice to Values: How to Speak Your Mind When You Know What's Right*. New Haven, CT: Yale University Press.
21. Horne, E., Pippin, S. E., & Vreeland, J. (2022). Examining the ethics education requirements for CPA candidates: State-by-state variations and the quest for mobility. *The CPA Journal*, *92*(3/4), 16–23.
22. Hurtt, R. K. (2010). Development of a scale to measure professional skepticism. *Auditing: A Journal of Practice & Theory*, *29*(1), 149–171.
23. Hurtt, R. K., & Thomas, C. W. (2008). Implementing a required ethics class for students in accounting: The Texas experience. *Issues in Accounting Education*, *23*(1), 31–51.

24. IAESB (International Accounting Education Standards Board). (2015). *Ethics Education Toolkit*. International Federation of Accountants.
25. IESBA. (2024). *Handbook of the International Code of Ethics for Professional Accountants (including International Independence Standards)*. International Ethics Standards Board for Accountants.
26. IFAC. (2021). *International Education Standard 4: Professional Values, Ethics, and Attitudes (Revised)*. International Federation of Accountants.
27. Jagger, S., Siala, H., & Sloan, D. (2016). It's all in the game: A 3D learning model for business ethics. *Journal of Business Ethics*, *137*(2), 383–403. https://doi.org/10.1007/s10551-015-2557-9
28. Jones, T. M. (1991). Ethical decision making by individuals in organizations: An issue-contingent model. *Academy of Management Review*, *16*(2), 366–395.
29. Klimek, J. L., & Wenell, K. J. (2011). Ethics in accounting: An indispensable course? *Academy of Educational Leadership Journal*, *15*(4), 107–118.
30. Lawrence, K. E., Reed, K. L., & Locander, W. (2011). Experiencing and measuring the unteachable: Achieving AACSB learning assurance requirements in business ethics. *Journal of Education for Business*, *86*(2), 92–99.
31. LCME (Liaison Committee on Medical Education). (2024). *Standards for Accreditation of Medical Education Programs Leading to the MD Degree (Standard 7.7: Medical ethics)*.
32. Liu, C., Yao, L. J., & Hu, N. (2012). Improving ethics education in accounting: Lessons from medicine and law. *Issues in Accounting Education*, *27*(3), 671–690.
33. Loeb, S. E. (1988). Teaching students accounting ethics: Some crucial issues. *Issues in Accounting Education*, *3*(2), 316–329.
34. Loeb, S. E. (2015). Active learning: An advantageous yet challenging approach to accounting ethics instruction. *Journal of Business Ethics*, *127*(1), 221–230.
35. Madison, R. L., & Schmidt, J. J. (2006). Survey of time devoted to ethics in accountancy programs in North American colleges and universities. *Issues in Accounting Education*, *21*(2), 99–109.
36. Manly, T. S., & Ritsema, C. M. (2018). A new approach to ethics education: Giving Voice to Values. *The Tax Adviser*, *49*(2), 118–120.
37. Marnburg, E. (2001). The questionable use of moral development theory in studies of business ethics: Discussion and empirical findings. *Journal of Business Ethics*, *32*, 275–283. https://doi.org/10.1023/A:1010643309056

38. Martinov-Bennie, N., & Mladenovic, R. (2015). Investigation of the impact of an ethical framework and an integrated ethics education on accounting students' ethical sensitivity and judgment. *Journal of Business Ethics*, *127*(1), 189–203. https://doi.org/10.1007/s10551-013-2007-5
39. Mayhew, B. W., & Murphy, S. H. (2009). The impact of ethics education on reporting behavior. *Journal of Business Ethics*, *86*(3), 397–416.
40. McManus, L., & Subramaniam, N. (2009). Ethical evaluations and behavioural intentions of early career accountants: The impact of mentors, peers, and individual attributes. *Accounting & Finance*, *49*(3), 619–643.
41. McPhail, K. (2001). The other objective of ethics education: Re-humanising the accounting profession—A study of ethics education in law, engineering, medicine and accountancy. *Journal of Business Ethics*, *34*(3–4), 279–298.
42. Mintz, S. M. (2019). A new approach to teaching ethical decision making to accounting students. *The CPA Journal*, *89*(9), 36–41.
43. Mintz, S. M. (2021). Accounting ethics education: A 43-year retrospective. *The CPA Journal*, *91*(8/9), 24–31.
44. NCBE (National Conference of Bar Examiners). (2025). *Multistate Professional Responsibility Examination (MPRE): About the MPRE*. https://www.ncbex.org/exams/mpre/
45. Nguyen, L. A., & Dellaportas, S. (2020). Accounting ethics education research: A historical review of the literature. In M. M. Pinheiro & A. J. Costa (Eds.), *Accounting ethics education: Teaching virtues and values* (pp. 44–80). Routledge. https://doi.org/10.4324/9780429321597-4
46. Nolder, C. J., & Kadous, K. (2018). Grounding the professional skepticism construct in mindset and attitude theory: A way forward. *Accounting, Organizations and Society*, *67*, 1–14. https://doi.org/10.1016/j.aos.2018.03.010
47. O'Leary, C., & Mohamad, S. (2006). The ethics of final year accountancy students: A tri-national study. *Malaysian Accounting Review*, *5*(1), 139–163.
48. Ponemon, L. A. (1993). Can ethics be taught in accounting? *Journal of Accounting Education*, *11*(2), 185–209.
49. Rest, J. R., & Narváez, D. (Eds.). (1994). *Moral Development in the Professions: Psychology and Applied Ethics*. Hillsdale, NJ: Lawrence Erlbaum.
50. Rexeisen, R. J., & Al-Khatib, J. (2009). Assurance of learning and study abroad: A case study. *Journal of Teaching in International*

Business, *20*(3), 192–207. https://doi.org/10.1080/08975930903099077

51. Robinson, C., Elson, R. J., & Williams, J. L. (2020). The current state of ethics education in accounting programs. *Global Journal of Accounting and Finance*, *4*(1), 111–128.
52. Robinson, S. N., Curtis, M. B., & Robertson, J. C. (2018). Disentangling the trait and state components of professional skepticism: Specifying a process of state scale development. *Auditing: A Journal of Practice & Theory*, *37*(1), 215–235.
53. Sherif, K., & Mohsin, H. (2021). The effect of emergent technologies on accountants' ethical blindness. *International Journal of Digital Accounting Research*, *21*, 61–94. https://doi.org/10.4192/1577-8517-v21_3
54. Thoma, S. J., & Dong, Y. (2014). The defining issues test of moral judgment development. *Behavioral Development Bulletin*, *19*(3), 55–61.
55. Thomas, S. (2012). Ethics and accounting education. *Issues in Accounting Education*, *27*(2), 399–418.
56. Thorne, L. (1998). The role of virtue in auditors' ethical decision making. *Research on Accounting Ethics*, *4*, 291–308.
57. Thorne LeClair, D., & Ferrell, L. (2000). Innovation in experiential business ethics training. *Teaching Business Ethics*, *4*(4), 361–377.
58. Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., ... & Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

Appendix A. Review Protocol, Evidence Matrix, Measures, and E⁴P Mapping

A1. Search Protocol and Selection (PRISMA-style summary)

The integrative scoping review covered the period from January 2000 through July 2025 across Scopus, Web of Science, ERIC, and Google Scholar. Search strings combined accounting with (ethics OR professional responsibility OR integrity) and (education OR curriculum OR pedagogy OR assessment). Peer-reviewed, English-language sources were included when they focused on accounting students or early-career accountants and reported learning or behavioral outcomes (e.g., moral awareness/sensitivity, moral reasoning, ethical intent/behavior, professional skepticism). Conceptual and standards documents (AACSB; IFAC IES 4; IESBA Code) were retained to contextualize expectations (AACSB, 2020, 2025; IFAC, 2021; IESBA, 2024).

Table A: PRISMA. Screening flow

Stage	***Count***
Records identified (Scopus=326; Web of Science=284; ERIC=97; Google Scholar initial pass=200)	907
Duplicates removed	186
Titles/abstracts screened	721
Full texts assessed for eligibility	146
Full texts excluded (not accounting=23; not educational intervention/outcome=41; insufficient outcome detail=19; commentary only=13)	96
Studies included in the qualitative synthesis	50
— of which: empirical (students=28; early-career/professional=9); conceptual/standards=13	50

Inclusion criteria: accounting-focused populations; educational or assessment interventions; outcomes related to awareness/sensitivity, reasoning/judgment, intent/behavior, or skepticism; sufficient methodological detail to extract design and measures. Exclusion criteria: non-accounting contexts without transfer rationale; opinion/commentary without outcomes; articles lacking relevant outcome measures. Counts reflect the finalized search; future updates can revise values without altering the synthesis.

A2. Evidence Matrix (selected sources informing mechanisms, pedagogy, and assessment)

Table A1: Evidence matrix (selected studies and standards)

Citation	*Setting / Design*	*Participants*	*Pedagogy / Instrument*	*Outcomes & Measures*	*Main finding(s)*	*Durability / Boundary notes*	*E⁴P link*
Dellaportas (2006)	Dedicated ethics course; quasi-experimental	Undergrads	Dilemma discussion; case analysis	Moral reasoning (DIT-style)	Significant gains in principled reasoning after course exposure	Durability not tested beyond the term	Exposure→ Engagement
Thomas (2012)	Scholarly review (auditing/ethics decision-making)	—	Synthesis of case-based and decision research	Ethical judgment/decision quality	Structured analysis and reflective discussion improve judgment; limits noted	Calls for stronger links to behavior	Engagement
Gentile (2010)	Pedagogical monograph	Business students/professionals	Scripted, rehearsed enactments under realistic constraints (GVV)	Implementation confidence; ethical intent	Rehearsal and scripting improve the ability to act on values	Transfer claims strong; contextualization advised	Enactment
Hurtt (2010)	Scale development and validation	Students / CPAs	Hurtt Professional Skepticism Scale (1–6)	Skepticism facets/trait	Validated instrument for audit-relevant skepticism	Trait-like; short-term shifts modest	Exposure/ Engagement
Robinson, Elson, and Williams	Program survey/landscape	U.S. accounting programs	Curriculum inventory (standalone vs	Coverage patterns; barriers	Growth in standalone courses;	Organizational constraints salient	Exposure (design)

(2020)			integrated)		variability and crowding persist		
Nguyen and Dellaportas (2020)	Historical review (ethics education)	—	Synthesis across pedagogies	Mapping of research streams	Recommends design-science and performance assessment	Calls for links to behavior and context	Engagement→ Enactment
Horne, Pippin, and Vreeland (2022)	Policy/requirements review	U.S. jurisdictions	State CPA ethics requirements	Structure & variation	Meaningful state-by-state variation; implications for mobility	Encourages clearer, portable outcomes	Partnership (policy)
AACSB (2020; 2025)	Standards	Business schools	Outcomes & societal impact	AoL expectations	Clear ethics outcomes and evidence required	Impact documentation emphasized	Exposure→ Partnership
IFAC (2021)	Education standard	Professional accountancy	Learning outcomes for ethics / skepticism	Competency benchmarks	Specifies values, public interest, and skepticism	Guides curriculum mapping	Exposure→ Engagement
IESBA (2024)	Ethics/Independence standard	Global profession	Principles-based code; threat-based approach	Principles & independence	Authoritative case anchor for instruction	Provides case scaffolds	Exposure

Synthesis: Awareness and reasoning respond to structured analysis and discussion (Dellaportas, 2006; Thomas, 2012), while enactment improves when students rehearse specific actions under constraint (Gentile, 2010). Skepticism shifts are typically smaller over short intervals (Hurtt, 2010). Program surveys and standards (Robinson et al., 2020; Horne et al., 2022; AACSB, 2020, 2025; IFAC, 2021; IESBA, 2024) support designs that make outcomes explicit and evidence visible. These patterns motivate the E⁴P mechanism (Rest and Narvaez, 1994; Jones, 1991).

A3. Measures, Scoring, and Assessment Artifacts (course-embedded pilot)

Moral awareness / Ethical sensitivity (O1). One AIS-specific vignette calibrated for moral-intensity cues is scored as issue recognition (proportion of key elements identified; 0–1 scale). Moral reasoning (O2). A DIT-style instrument yields a summary reasoning score. Professional skepticism (O3). The Hurtt (2010) scale (1–6) captures skepticism relevant to audit/control judgments. Implementation confidence / Enactment (O4). A GVV performance task requires students to script and rehearse values-consistent responses under realistic constraints; a four-dimensional rubric scores awareness, reasoning, voice/courage, and outcome on a 1–4 scale.

Table A2: Measures and scoring summary

Outcome	***Operationalization***	***Scale***	***Primary Evidence / Rationale***
O1 Ethical Sensitivity	AIS vignette with moral-intensity cues; issue recognition proportion	0–1	Structured analysis improves recognition (Thomas, 2012)
O2 Moral Reasoning	DIT-style summary (P/N2 proxy)	Scale score	Dedicated instruction linked to gains (Dellaportas, 2006)
O3 Professional Skepticism	Hurtt Professional Skepticism Scale	1–6	Validated instrument; shifts modestly (Hurtt, 2010)
O4 GVV Enactment	GVV script & rehearsal; 4-dimension rubric (A/R/V/O)	1–4	Rehearsal strengthens implementation (Gentile, 2010)

Mastery thresholds were set a priori at O1 ≥ .80 (high sensitivity) and O4 ≥ 3.25 (strong enactment). Artifacts for assurance include vignette scores, DIT-style summaries, GVV rubrics, and a brief internship mentor check-in (two ratings and one comment) when available. Artifacts align with AACSB outcomes and IES 4 competencies (AACSB, 2020, 2025; IFAC, 2021).

A4. Mapping of Practices to E⁴P and to Standards

Table A3: Practices → E⁴P → Standards → Artifacts

E⁴P stage	*Exemplary practices (adaptable)*	*Primary measures*	*Standards alignment*	*AoL artifact(s)*
Exposure	Introductory ethics seminar; standards primers embedded in AIS/Audit/Tax; case pre-reads	Awareness checklist; short quizzes	AACSB outcomes; IES 4 values/public interest; IESBA principles	Intro quiz; pre-read notes
Engagement	Structured case analysis; dilemma discussions; short simulations	Ethical sensitivity vignette; DIT-style reasoning	IES 4 reasoning; skepticism calibration	Vignette score; reasoning summary
Enactment	GVV scripting and rehearsal; role-play with realistic constraints	GVV rubric (A/R/V/O)	IES 4 behavioral demonstration; IESBA principles-in-action	GVV rubric + reflection
Partnership	Internship mentor check-in; early-career follow-up; employer roundtables	Brief ratings + open comment	AACSB outcomes/impact	Mentor form; aggregate summary

A5. Risk-of-Bias and Evidence Limitations

The synthesis is limited to English-language sources and relies on widely used educational outcomes (e.g., DIT-style reasoning, ethical sensitivity tasks, GVV performance), imperfect proxies for real-world behavior. Many studies report short-term changes without long-term follow-up; results are likely influenced by contextual moderators such as ethical climate, incentives, and moral-intensity cues. These constraints motivate the emphasis on enactment tasks, practice-linked artifacts, and internship follow-ups to assess durability (Rest and Narvaez, 1994; Jones, 1991).