Journal of International Education in Business

Replicating the impact of gamified learning on awareness of management competencies: a two-cohort mixed-methods study

Gelmar García-Vidal
Department of Management Science, Universidad UTE, Quito, Ecuador

Antonio J. Canela-Ruano
Methodological Advisory Unit, Universidad UTE, Quito, Ecuador

Laritza Guzmán-Vilar
Independent Researcher, Quito, Ecuador, and

Reyner Pérez-Campdesuñer
Department of Management Science, Universidad UTE, Quito, Ecuador

Received 12 February 2026
Revised 29 May 2026
28 June 2026
Accepted 13 July 2026

Abstract

Purpose – This study aims to examine whether a brief simulation game produces changes in management students' awareness of managerial competencies, distinguishing awareness revelation from skill development and investigates temporal pressure as the underlying mechanism.

Design/methodology/approach – A convergent mixed-methods design combined a quasi-experimental pre-post study across two independent cohorts ($n = 91$) with semi-structured interviews ($n = 24$). The 90-minute Shipwreck survival simulation game – individual prioritization, team consensus under a 15-minute constraint, strategic decision and structured debriefing – was assessed across three awareness dimensions: valorization, self-awareness and applied comprehension. Interview themes were independently double-coded (Cohen's $\kappa = 0.73$).

Findings – Significant pre-post improvements emerged across all dimensions ($d = 2.81$–3.85, $p < 0.001$), largest for applied comprehension and smallest for valorization (ceiling effect). Pressure-management subfactors showed the strongest effects ($d \approx 4.16$–4.56) and the pattern was replicated across cohorts. These magnitudes are unusually large for educational interventions and should be interpreted considering the within-subjects design and self-report measurement. Qualitative accounts converged with these results, identifying the transition from unconscious to conscious incompetence as the core learning process.

Research limitations/implications – Self-report measures and the absence of a control group constrain causal inference; single-institution sampling limits generalization. Replication logic and the theoretical coherence of the differential effects support their interpretation as awareness revelation.

Practical implications – Simulation games function as diagnostic interventions optimally positioned before skill-building modules, using temporal pressure to expose competency gaps and debriefing to consolidate metacognitive awareness.

Journal of International Education in Business
Emerald Publishing Limited
2046-469X
DOI 10.1108/JIEB-02-2026-0027

Originality/value – The study distinguishes competency awareness from behavioral competence as distinct outcomes, identifies a differential pattern consistent with temporal pressure as a candidate revelatory mechanism and extends the evidence base to under-researched Latin American management education.

Keywords Gamification, Experiential learning, Competency awareness, Management education, Mixed methods

Paper type Research paper

Introduction

Management education faces a persistent tension between teaching *about* management and developing capacity *for* management (Mintzberg, 2004), a tension that remains central to current scholarship on business education. Recent reviews argue that the narrow forms of management education currently adopted by most business and management schools cannot produce the graduate competences needed for contemporary practice (Mason, Anderson, Black and Roberts, 2024) and field studies continue to document a growing disconnect between the knowledge taught in business school and the needs of management practice (Liu, Lv, Huang and Wang, 2024). Traditional pedagogies, case analysis, lectures, structured exercises, excel at transmitting conceptual knowledge yet frequently fail to convey the experiential reality of managerial work: decision-making under uncertainty, coordination under time pressure and the gap between theoretical prescriptions and situated practice (Pettigrew and Starkey, 2016). Experiential learning methods, including simulations and games, have emerged as complementary approaches that compress managerial complexity into bounded, reflective experiences (Kolb, 2015). However, the specific mechanisms through which these experiences produce learning remain insufficiently specified and their outcomes are often conflated under undifferentiated constructs of "engagement" or "skill development." Gamification, the integration of game design elements into educational contexts, has gained substantial traction as a mechanism for addressing this tension (Deterding, Dixon, Khaled and Nacke, 2011; Legaki, Karpouzis, Assimakopoulos and Hamari, 2021). Yet the term encompasses heterogeneous implementations, from superficial point systems to immersive simulation games (Silva, Gouveia-Rodrigues and Leal, 2019; Triantafyllou, GAdis and Sapounidis, 2025). Survival simulation games, exemplified by the "Shipwreck" or "Lost at Sea" exercise, represent a distinct subset: structured scenarios where teams prioritize resources under temporal and environmental constraints, receiving immediate feedback against expert benchmarks (Johnson and Johnson, 1999). These exercises exhibit core ludic properties, narrative immersion, formal rules, quantifiable outcomes and the "magic circle" of suspended real-world consequences (McGonigal, 2011; Salen and Zimmerman, 2003) yet have been predominantly deployed as team-building diagnostics rather than systematically examined as vehicles for competency development.

Recent scholarship suggests that simulation games may produce effects distinct from skill acquisition: they may function as "revelatory" experiences that render visible the gap between assumed and actual capabilities (Crookall, 2010; Legaki *et al.*, 2021). Under conditions of temporal pressure, learners confront the limitations of their heuristic reasoning and the complexity of executing theoretical knowledge under realistic constraints (Durrani, Al-Naymat, Ayoubi, Kamal and Hussain, 2022; Memar, Sundström and Larsson, 2021). This positions competency *awareness* – recognition of what one does not know or cannot do – as a valid outcome distinct from behavioral competence itself. Such awareness aligns with Gordon and Burch (1974) stage of "conscious incompetence" a necessary precursor to targeted skill development and with Mezirow (1991) "disorienting dilemmas" that trigger perspective transformation.

Empirical examination of this mechanism remains scarce. Studies of gamification in management education have predominantly assessed motivation, engagement or declarative learning outcomes (Armenia, Barnabè, Nonino and Pompei, 2024; Gatti, Ulrich and Seele, 2019), while research on survival simulations has focused on team processes or decision quality (Faria, 2001). The specific contribution of the simulation *game* frame – particularly its combination of narrative immersion, temporal pressure and post-experience debriefing – to awareness of managerial competencies has not been isolated. Moreover, existing evidence derives predominantly from North American and European contexts, limiting generalization to educational systems where experiential methods remain novel (Castillo-Parra, Hidalgo-Cajo, Vásconez-Barrera and Oleas-López, 2022).

This study addresses these gaps through a two-cohort quasi-experimental investigation of the Shipwreck simulation in an introductory management course at a private Ecuadorian university, examining pre-post changes across three awareness dimensions: valorization (perceived importance of managerial competencies), self-awareness (recognition of capability gaps) and applied comprehension (understanding of contextual application). Three research questions guide the inquiry:

RQ1. To what extent does a brief survival simulation produce changes in students' awareness of managerial competencies, valorization, self-awareness and applied comprehension, in a first-year, emerging-market university context?

RQ2. Is there a systematic differential pattern across the three awareness dimensions consistent with temporal pressure as a specific revelatory mechanism and does this pattern replicate across independent cohorts?

RQ3. Which experiential processes, as reported by students, account for the observed quantitative changes and how do these processes relate to the diagnostic, rather than skill-developing, nature of the simulation?

The questions are addressed sequentially: *RQ1* through pre-post magnitude analysis across dimensions, *RQ2* through subfactor and cross-cohort pattern matching and *RQ3* through thematic analysis of post-intervention interviews integrated with the quantitative findings.

The contribution is threefold. First, we distinguish awareness revelation from skill development as distinct pedagogical outcomes, advancing simulation game conceptualization in management education. Second, we document a differential pattern consistent with temporal pressure as a candidate mechanism for exposing competency gaps, informing experiential intervention design. Third, we extend the gamified learning evidence base to Latin American higher education, where experiential methods' novelty may intensify awareness effects. The study suggests that brief simulations may produce awareness of competency application complexity, a necessary foundation for skill-building, rather than developing competencies themselves.

Literature review

Gamification, the use of game design elements in non-game contexts, has proliferated in management education under divergent operationalizations (Deterding *et al.*, 2011). Early approaches emphasized extrinsic motivators such as points, badges and leaderboards added to conventional instruction (Hamari, Koivisto and Sarsa, 2014; Zichermann and Cunningham, 2011), whereas more recent scholarship distinguishes "gamification" from "serious games" and "simulation games," the latter embedding game mechanics structurally rather than superimposing them (Silva *et al.*, 2019; Triantafyllou *et al.*, 2025). This distinction carries pedagogical consequences: simulation games activate what Salen and Zimmerman (2003)

term the "magic circle", a bounded space where real-world consequences are suspended yet engagement is heightened, potentially generating distinct learning mechanisms from surface-level gamification (Almeida and Simões, 2019; López, Arias-Oliva, Pelegrín-Borondo and Marín-Vinuesa, 2021).

Survival scenarios exemplify the simulation game form. The "Shipwreck" exercise and its variants have been deployed in management curricula since the 1970s yet are predominantly framed as team-building diagnostics or icebreakers rather than as structured games (Johnson and Johnson, 1999). This framing has obscured their inherent ludic properties: narrative immersion ("you are aboard a sinking vessel"), formal rules (time constraints, resource scarcity, quantifiable outcomes) and immediate feedback against expert benchmarks (Kapp, 2012). These properties align with McGonigal (2011) characterization of "urgent optimism", the psychological state induced by challenging yet achievable objectives under time pressure and with Crookall (2010) emphasis on debriefing as the mechanism converting game experience into transferable learning. Recent empirical evidence supports this reconceptualization: game-and simulation-based approaches have outperformed traditional instruction in project management (Armenia *et al.*, 2024), enhanced practical skill acquisition in portfolio management (Barbosa and de Ávila Rodrigues, 2020) and improved problem-solving in entrepreneurship education (Crespo-Martínez, Bueno and Gallego, 2025).

The theoretical implications remain underdeveloped. While gamification research has examined motivation and engagement (Sailer, Hense, Mayr and Mandl, 2017; Yalcin, Pehlivan and Nistor, 2020) and simulation research has assessed decision-making outcomes (Faria, 2001; Gatti *et al.*, 2019), how simulation games specifically affect awareness of managerial competencies has received limited attention. No study has isolated whether the "game against nature" structure (Rapoport, 1966), where teams confront environmental uncertainty rather than competing directly – produces differential effects on self-awareness versus skill acquisition. This distinction matters: if simulation games primarily reveal competency gaps rather than develop them, their curricular positioning shifts from capstone to foundational diagnostic interventions (Hyams-Ssekasi and Taheri, 2022; Jerônimo Soares, Neves Bezerra de Melo, de Tarso Alves Dantas, da Silva and Genuino, 2024).

Temporal pressure as revelatory mechanism

The active mechanisms within simulation games require specification. Temporal pressure, a central design element in survival scenarios, has been treated inconsistently in the literature, either as a confound to be minimized (Kolb, 2015) or as a stressor impairing performance (Maule, Hockey and Bdzola, 2000). An alternative conceptualization, derived from dual-process theory (Grayot, Beck and Heijmeskamp, 2024), suggests that time constraints may function as "revelatory" rather than merely "impairing": by forcing reliance on heuristics under cognitive load, pressured contexts expose the gap between declarative knowledge and procedural capability (Gordon and Burch, 1974). The simulation game frame, where failure carries no external penalty, may render such exposure psychologically safe, accelerating the transition from unconscious to conscious incompetence.

Recent empirical work supports this mechanism: gamified forecasting education improved learning about heuristics and biases (Legaki *et al.*, 2021), long-term gamified assessment amplified engagement and awareness of learning gaps (Montero-Benavides, Albort-Morant and Masero-Moreno, 2025) and gamified experiential learning prompted causation and effectuation behaviors under scale and time pressure (Memar *et al.*, 2021). However, these studies focused predominantly on skill development or engagement rather than on the specific awareness shifts, valorization, self-awareness and applied comprehension, that may constitute the primary contribution of brief, intensive simulation games.

Awareness as a managerial competency outcome

Beyond skill development, management education literature increasingly recognizes awareness as a substantive pedagogical outcome in its own right (Eurich, 2018). Awareness, in the present study, refers to the metacognitive recognition of what one knows, does not know and is unable to do in a given competency domain, a construct distinct from declarative knowledge (knowing about), procedural skill (knowing how) and motivated valuing (caring about). This conceptualization draws on the metacognitive tradition (Flavell, 1979; Schraw and Moshman, 1995), which distinguishes knowledge of cognition from regulation of cognition and on competence-development models that situate awareness as the precondition for targeted skill acquisition (Gordon and Burch, 1974). Within international business education, where managerial work spans cross-cultural decision contexts and pressure-laden coordination demands, the diagnostic recognition of one's competency boundaries is increasingly treated as foundational to managerial development rather than incidental to it (Bedwell, Fiore and Salas, 2014; Kothiyal, Bell and Clarke, 2018).

Three dimensions operationalize this outcome. Valorization – the perceived importance students attribute to a set of managerial competencies – rests on the attitudinal literature relating value judgments to professional identity formation (Ajzen, 1991; Ashforth, Harrison and Corley, 2008) and on the management-education tradition that treats the legitimation of competencies as the affective foundation of subsequent skill investment (Boyatzis, 1982; Mintzberg, 2004). Self-awareness corresponds to students' recognition of personal capability gaps, anchored in the organizational-behavior literature on managerial self-awareness as a predictor of leader effectiveness (Eurich, 2018; Sutton, Williams and Allinson, 2015) and in the conscious incompetence stage of competence development (Gordon and Burch, 1974), where acknowledging deficiency is itself the developmental outcome. Applied comprehension refers to the contextual understanding of when and how competencies operate in practice, drawing on the situated and reflective learning traditions that frame managerial knowing as activity-bound rather than as the abstract possession of principles (Schön, 1983).

Temporal pressure occupies a distinctive position within this conceptual frame. Beyond its role as a stressor that impairs performance (Maule *et al.*, 2000), pressure functions as a competency-disclosing condition: under cognitive load, learners default to heuristic reasoning and become observable, to themselves and to others, in their default decision patterns (Edland and Svenson, 1993; Kahneman, 2011). In bounded pedagogical settings, this disclosure occurs without the consequential cost of real-world managerial error, generating the disorienting dilemma that Mezirow (1991) identified as the activating condition of transformative learning. Awareness, in this sense, is not a residual of compressed experiential learning but its primary intended product, particularly in introductory management education in international contexts where students' first encounter with experiential pedagogy is itself the disrupting moment.

The emerging market educational context

Latin American management education has historically emphasized magisterial instruction over experiential methods (Castillo-Parra *et al.*, 2022; Davila and Elvira, 2012). Consequently, student populations in this context may exhibit lower baseline familiarity with simulation games, potentially amplifying awareness effects as novel disruption. The absence of empirical evidence from these contexts limits generalization of existing gamification findings, which derive predominantly from North American and European higher education institutions (Jerônimo Soares *et al.*, 2024; Silva *et al.*, 2019).

Recent international studies confirm that game-based learning effects are contextually contingent: gamification enhanced engagement with nuanced effects on skill development in the United Arab Emirates (Durrani *et al.*, 2022), competitive e-commerce games developed

specific entrepreneurial competencies Ho and Chen (2023) and serious games for business ethics elicited significant cross-cultural differences across British universities Siala *et al.* (2025) Yet none has examined survival simulation games in Latin American introductory management education, where the novelty of experiential methods may intensify awareness effects.

Taken together, these gaps call for empirical examination of simulation games as distinct pedagogical forms, with attention to:

- their effects on competency awareness versus behavioral competence;
- temporal pressure as a revelatory mechanism; and
- the robustness of effects across contexts with varying experiential learning traditions.

The present study addresses these gaps through a two-cohort replication of the Shipwreck simulation game in an Ecuadorian business program.

Method

Design and context

This study used a quasi-experimental pre-post design with two independent cohorts to examine the effects of a structured experiential learning intervention on first-year students' awareness of managerial competencies. The investigation originated in the reflective teaching practice of the instructor-researcher and is therefore situated within the tradition of educational action research, in which the systematic study of one's own classroom serves both pedagogical improvement and the generation of transferable knowledge (Coghlan and Jacobs, 2005; Starr-Glass, 2011). Consistent with this tradition, the second cohort functioned as a deliberate replication of the first: rather than modifying the intervention between cohorts, the study held the protocol constant to test whether the learning effects observed in the first delivery would recur under independent sampling, thereby combining the practitioner orientation of action research with the inferential logic of replication (Coghlan, 2010). The research was conducted at a private university in Ecuador, within the introductory course General Management. The intervention, the "Shipwreck" survival simulation, was positioned during the initial weeks of the academic semester (Sessions 3–4), preceding formal instruction in organizational behavior, leadership and team dynamics, so that measured awareness reflected students' pre-instructional state rather than recently taught content.

The design embedded measurement within authentic curricular practice, enhancing ecological validity. Two consecutive cohorts (2024–2, $n = 43$; 2025–1, $n = 48$) completed identical protocols. Because the simulation was a required curricular activity, a no-treatment control group was neither pedagogically nor ethically feasible: withholding a graded learning activity from a subset of enrolled students would have disadvantaged them academically. The two-cohort replication design was adopted as the methodologically and ethically appropriate alternative, providing a form of analytical robustness in the absence of a control condition. Participants were first-year Administration students (18–20 years old; 58–64% female across cohorts) enrolled in General Management (Course Code 1136), the foundational course of the International Business program, none of whom had prior exposure to formal management education or experiential learning methodologies.

Ethical procedures. The study was conducted in accordance with institutional research-ethics requirements and the principles of the Declaration of Helsinki. Although participation in the simulation itself was a mandatory graded component of the course (25% of the final grade), participation in the research, completion of the pre- and post-test surveys and the

interviews for analytic purposes, was entirely voluntary and independent of course assessment. Prior to the pretest, all students received a written information sheet describing the study's purpose, the voluntary nature of research participation, the absence of any academic consequence for declining, the anonymization of responses through coded identifiers and the secure storage of data accessible only to the research team. Written informed consent was obtained on this basis. Four students declined research participation; their survey data were excluded from all analyses, although they completed the simulation as part of normal coursework. Interview participants provided separate written consent for audio recording and all transcripts were pseudonymized prior to analysis.

The intervention: the shipwreck simulation

The intervention consisted of a 90-minute structured simulation based on the *Lost at Sea* consensus-seeking task originally developed by Nemiroff and Pasmore (1975) (see also Nemiroff, Pasmore and Ford (1976)), one of the most enduring and widely documented survival-scenario exercises in management and organizational behavior education. The exercise was selected for three reasons. First, its survival-decision structure under resource scarcity and time constraint maps directly onto the managerial competencies of interest, decision-making under uncertainty, coordination and pressure tolerance, making it substantively appropriate for an introductory management course rather than a generic team-building filler. Second, it possesses an established reference ranking that enables an objective, replicable performance metric across cohorts, a property that purely open-ended group tasks lack and that is essential for the cross-cohort replication design adopted here. Third, its brevity and logistical simplicity allow full implementation, including individual ranking, team consensus and structured debriefing, within a single 90-minute class session, fitting the constraints of authentic curricular delivery. The pedagogical framing followed cooperative-learning principles (Johnson and Johnson, 1999), aligning with the course's emphasis on gamification and cooperative learning and the exercise supported the course Learning Outcome of analyzing the organization across the phases of the administrative process by operationalizing planning, coordination and control functions under compressed temporal conditions.

Students were randomly assigned to teams of 6–8 members and presented with the standardized maritime emergency of the Lost at Sea task: a fishing vessel sinking in the South Pacific, approximately 1,000 nautical miles from the nearest shore, with 15 salvaged items (sextant, shaving mirror, 20-liter water container, mosquito netting, case of food rations, Pacific Ocean map, air cushion, 8-liter oil-petrol can, transistor radio, shark-protection device, 6 m^2 of transparent nylon sheeting, two bottles of rum, 4.5-meter nylon rope, two boxes of chocolate, fishing rod), along with a rubber life raft and oars, requiring prioritization for survival. The reference ranking against which performance was scored is the one established in the original task, derived from maritime survival criteria attributed to the USA Coast Guard, which prioritizes items that attract rescue (signaling devices) and sustain life until rescue arrives over items supporting navigation or escape (Nemiroff and Pasmore, 1975). Although this reference ranking provides a defensible benchmark for scoring deviation, the task does not admit a single incontrovertible solution; performance scores therefore index relative distance from an established expert criterion rather than absolute correctness and the pedagogical value resides primarily in the collective sense-making process the task elicits rather than in the attainment of a "correct" answer.

The protocol comprised four phases designed to enact Kolb (2015) experiential learning cycle within a compressed pedagogical window, as summarized in Table 1.

The sequence moved from individual judgment under time pressure (concrete experience), through unstructured team deliberation that surfaced emergent leadership and

Table 1. Four-Phase Structure of the shipwreck simulation protocol

Phase	Duration	Activity	Cognitive demand induced	Kolb (2015) cycle stage	Data generated
1. Individual ranking	15 min	Each participant independently prioritizes the 15 items under time pressure	Individual judgment under uncertainty and time constraint	Concrete experience	Individual deviation scores versus reference ranking
2. Team consensus	15 min	Teams reach a unified ranking through unstructured deliberation without facilitation	Collective sense-making; emergent leadership and conflict management	Reflective observation	Team consensus score; synergy index
3. Strategic decision	5 min	Teams choose between rowing toward shore or awaiting rescue	Decision-making under irreducible uncertainty	Abstract conceptualization	Strategic choice (row/wait)
4. Debriefing	55 min	Facilitated "4F" reflection (Facts, Feelings, Findings, Future) redirecting attention from outcome to process	Metacognitive reflection on decision processes and group dynamics	Active experimentation	Qualitative reflection feeding interview themes

Note(s): Total session duration ≈ 90 min. The reference ranking derives from the original task (Nemiroff and Pasmore, 1975). Instructors followed identical scripts for phase transitions and debriefing prompts across both cohorts to ensure implementation fidelity
Source(s): Authors' own work

conflict-management patterns (reflective observation), to an explicit strategic decision under irreducible uncertainty (abstract conceptualization) and finally to a structured "4F" debriefing (Facts, Feelings, Findings, Future) that redirected attention from outcome to process (active experimentation). The 15 ranking items operated as material anchors for this cycle: each item generated a discrete decision occasion whose post-task examination during debriefing rendered visible the heuristics that had produced the ranking, thereby converting individual choices into objects of metacognitive reflection on managerial behavior. Instructors, trained in facilitation but blind to research hypotheses, standardized implementation across cohorts using identical scripts for phase transitions and debriefing prompts.

The 15-minute constraints were deliberately stringent to induce cognitive load and simulate realistic managerial pressure where decisions precede complete information acquisition. The simulation generated objective performance metrics: individual scores (deviations from expert rankings), team scores (consensus quality), synergy index (team score − mean individual score), strategic choice (row/wait) and survival threshold (≤60 points). These metrics characterized sample performance and contextualized self-reported awareness changes but were not analyzed inferentially.

Instrumentation

Awareness of managerial competencies was measured with a purpose-built 24-item self-report instrument administered immediately before and after the intervention. The instrument operationalizes three theoretically derived dimensions designed to capture awareness as a distinct outcome from behavioral competence (Eurich, 2018; Gordon and Burch, 1974). Each dimension is measured with 8 items rated on a 5-point Likert scale (1 = *strongly disagree* to 5 = *strongly agree*), yielding subscale scores between 8 and 40 and a total score between 24 and 120.

Dimension 1: Valorization assesses the perceived importance students attribute to eight managerial competencies – influence without formal authority, initiative under uncertainty, argumentation, active listening, leveraging collective intelligence, negotiation, pressure tolerance and coordination – following operationalizations consistent with the managerial competency literature (Boyatzis, 1982; Mintzberg, 2004). *Dimension 2: Self-awareness* measures students' recognition of personal capability gaps across the same eight domains, using items framed in the direction of perceived need for improvement (e.g. "I need to improve my ability to influence group decisions"). Because items are written in the direction of acknowledged need, higher scores directly index greater self-recognition of developmental gaps; no reverse coding is applied. This scoring convention is theoretically grounded in the *conscious incompetence* stage of the competence development model (Gordon and Burch, 1974), in which awareness of one's own limitations is itself the developmental outcome and is expected to increase, rather than decrease, after a revelatory pedagogical experience. *Dimension 3: Applied comprehension* assesses students' understanding of the contextual application of managerial competencies – the situational *when* and the procedural *how* of competency deployment – operationalized following the practical-knowledge literature on managerial situational judgment (Schön, 1983).

Instrument development and validation

The instrument was developed in Spanish through a four-stage procedure:

(1) item generation derived from the three theoretical dimensions and the eight competency domains;

(2) content validity review by a panel of three management-education experts who evaluated relevance, clarity and theoretical fit;

(3) cognitive pretesting with a separate cohort ($n = 32$) to assess comprehensibility, semantic equivalence across managerial constructs and time to completion, with minor wording adjustments made for items that the pretest cohort flagged as ambiguous; and

(4) confirmation of factorial structure and internal consistency on the main study sample ($N = 91$) reported below. Construct validity was examined through an exploratory factor analysis on the pretest data of the main sample ($N = 91$, 24 items).

Sampling adequacy was satisfactory (KMO = 0.79; item-level KMO range = 0.72–0.89) and Bartlett's test of sphericity was significant, $\chi^2(276) = 1092.96$, $p < 0.001$, supporting factorability. Principal axis extraction with Promax rotation yielded a three-factor solution – pre-specified on theoretical grounds and consistent with parallel-analysis evidence – accounting for 56.75% of the total variance.

All 24 items loaded on their theoretically assigned factor with primary loadings between 0.65 and 0.82, no cross-loading above 0.20 and communalities ranging from 0.43 to 0.68 ($M = 0.57$), indicating a clean simple structure (see online Appendix A for the full pattern matrix). Internal consistency was good across all dimensions and measurement occasions (Table 2): Valorization ($\alpha_{pre} = 0.89$, $\alpha_{post} = 0.86$), Self-Awareness ($\alpha_{pre} = 0.90$, $\alpha_{post} = 0.88$) and Applied Comprehension ($\alpha_{pre} = 0.87$, $\alpha_{post} = 0.86$).

Qualitative component

Following the post-test administration, semi-structured interviews were conducted with 24 students (12 per cohort) selected through purposive stratified sampling across the four performance quartiles and the three synergy bands (positive, neutral, negative) derived from the simulation outcomes. This stratification ensured representation of contrasting experiential trajectories, from teams whose collective decisions outperformed individual averages to teams where group process diminished individual judgment and thereby maximized variability in the experiential mechanisms accessible through interview data.

The interview protocol adopted a minimalist design, with a maximum duration of 10 min per participant, structured around four open-ended core questions probing:

(1) self-observed behavior during the simulation;

(2) perceived learning about managerial competencies;

(3) recognition of personal capability gaps; and

(4) anticipated professional application of the experience.

Table 2. Internal consistency reliability by dimension and measurement occasion ($n = 91$)

Dimension	Items	α Pre-test	α Post-test
D1. Originalidad/valorization	8	0.89	0.86
D2. Self-awareness	8	0.90	0.88
D3. Applied comprehension	8	0.87	0.86

Note(s): α = Cronbach's alpha. All coefficients meet the conventional threshold for good internal consistency (Nunnally and Bernstein, 1994)
Source(s): Authors' own work

Non-directive probing was used only to invite elaboration on initial responses, with no follow-up questions introducing new content. Interviews were audio-recorded with participants' written consent, transcribed *verbatim* in the original language (Spanish) and pseudonymized prior to analysis. Procedural documentation of the thematic analysis is reported in the *data analysis* subsection above.

Mixed-methods integration

The study adopted a *convergent parallel mixed-methods design* (Creswell and Plano Clark, 2018), in which quantitative and qualitative strands were collected during the same intervention window, analyzed independently using procedures appropriate to each data type and subsequently integrated at the interpretation stage to address complementary aspects of the research questions (see Figure 1).

The quantitative strand documents the *magnitude* and *pattern* of pre-post change across awareness dimensions and subfactors; the qualitative strand explicates the *experiential mechanisms* through which the intervention produced those changes. Integration is achieved through *joint display* (Guetterman, Fetters and Creswell, 2015): quantitative patterns are matched against thematic categories from the interviews, allowing convergence, expansion and divergence to be examined systematically. This design is appropriate when neither data strand alone can answer the research questions, quantitative data establish *whether* and *how much* awareness shifts, while qualitative data illuminate *why* and *through which experiential processes* and when the researcher has the resources to collect both strands within a bounded intervention (Tashakkori, Johnson and Teddlie, 2020).

Data analysis

Quantitative analysis. All analyses were conducted in R (version 4.3.1) using the *stats*, *psych* and *effect size* packages. Pre-post comparisons were tested with Wilcoxon signed-rank tests for paired samples. Although total dimension scores aggregate eight Likert items and approximate continuous measurement, three considerations justified a uniform non-parametric approach. First, Shapiro–Wilk tests indicated departures from normality in several pre-post distributions, particularly at post-test where ceiling effects compressed upper-range variance (e.g. D1 post-test: $W = 0.94$, $p = 0.03$; D3 post-test: $W = 0.93$, $p = 0.01$ in Cohort 2024–2). Second, the underlying item-level measurement is ordinal and although central limit considerations support parametric analysis of aggregated Likert scores in large samples, a conservative uniform non-parametric approach preserves consistency across all reported contrasts – dimensions, subfactors and total scores – without introducing decision rules that vary by variable. Third, parallel paired *t*-tests conducted as a sensitivity analysis (reported in Online Appendix B) confirmed convergence with the non-parametric results: all pre-post contrasts on dimensions and on the total scale were significant at $p < 0.001$ in both cohorts and the substantive dimensional hierarchy *D3 > D2 > D1* was preserved identically under both procedures, confirming that the choice of non-parametric versus parametric testing does not affect substantive inference.

Effect sizes. For each Wilcoxon test, the effect size r was computed as $r = |Z|/\sqrt{n}$, where n is the number of paired observations (Fritz, Morris and Richler, 2012; Rosenthal, 1991). For comparability with educational-intervention benchmarks expressed in standardized mean differences, r values were approximated to Cohen's d using $d \approx 2r/\sqrt{(1-r^2)}$ (Borenstein, Hedges, Higgins and Rothstein, 2009). Because the present design is within-subjects, this conversion yields d magnitudes that are systematically inflated relative to between-subjects equivalents (Lakens, 2013); converted values are therefore reported for interpretive

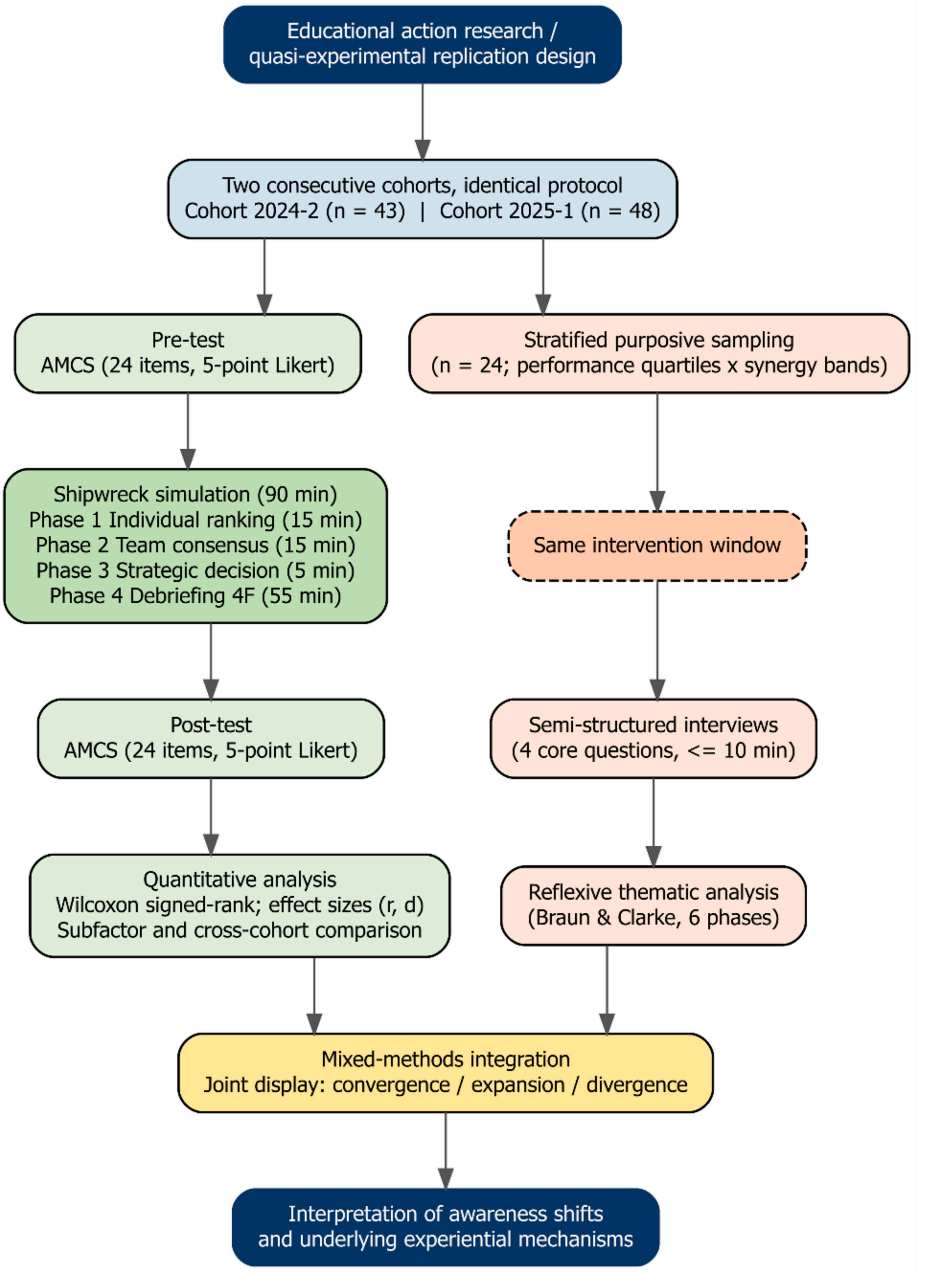

Figure 1. Methodological process of the convergent parallel mixed-methods design. The two cohorts followed an identical protocol, enabling the second cohort to function as a literal replication of the first. The quantitative and qualitative strands were collected within the same intervention window and analyzed independently using procedures appropriate to each data type, then integrated through a joint display at the interpretation stage. AMCS = awareness of managerial competencies scale; 4F = facts, feelings, findings, future debriefing framework

Source: Authors' own work

benchmarking only and are not directly comparable to *d* values from between-groups studies.

Replication and cohort convergence. The two cohorts were treated as independent replications of the same intervention. Cohort convergence was assessed through descriptive pattern matching across three analytic levels: the dimension-level ordering (D3 > D2 > D1), the subfactor-level hierarchy within each dimension and the magnitude of pre-post change. Convergence across these three patterns under independent sampling provides analytical robustness in the absence of an active control condition, following the literal-replication logic articulated by Yin (2018) for case-based designs. Inferential cross-cohort testing was not performed because the design does not treat cohort as an experimental factor; instead, cross-cohort variance is reported descriptively as a robustness indicator.

Operationalization of derived variables. Three derived variables are reported in subsequent sections. *Team synergy* is operationalized as the mean of individual error scores within the team minus the team consensus error score; a positive value (team error below the mean individual error) indicates positive synergy, a negative value indicates process loss. *Survival threshold* is operationalized as a categorical indicator distinguishing teams whose total error scores fell at or below 60 points from teams above this cutoff. *Strategic choice* is operationalized as a binary classification of the team's collective decision between rowing toward the shipping lane (active strategy) and remaining adrift (passive strategy), recorded by the facilitator from the post-task team reports.

Qualitative analysis. Interview transcripts were analyzed following reflexive thematic analysis (Braun and Clarke, 2006, 2019). Coding proceeded inductively from the *verbatim* transcripts through six iterative phases: familiarization with the data corpus, generation of initial descriptive codes, search for candidate themes, review of themes against extracts and across cases, definition and naming of final themes and integration with quantitative patterns. Themes were anchored to the four-question interview protocol but were not constrained by it; emergent themes that cut across questions were retained. Theme frequencies are reported as the number of participants (out of $n = 24$) whose accounts contained material coded under each theme.

To establish the trustworthiness of the thematic coding, two researchers coded all 24 transcripts independently. The primary researcher generated the initial codebook of six themes; a second coder, external to the original analysis, then applied the codebook to the full corpus without access to the first coder's decisions. Inter-coder agreement was assessed with Cohen's kappa computed per theme and pooled across all coding decisions. Pooled agreement was substantial [$\kappa = 0.73$; 88.2% raw agreement; Landis and Koch (1977)], with per-theme values ranging from $\kappa = 0.41$ to $\kappa = 0.90$. The two themes with lower coefficients – self-awareness and pressure management – reflected legitimate conceptual boundaries (self-directed versus group-directed insight; time-induced cognitive degradation versus physiological stress response) and, in the case of self-awareness, the high prevalence of the theme, which depresses kappa even at high raw agreement [the prevalence paradox; Cicchetti and Feinstein (1990)]. All 17 discrepancies were resolved through consensus discussion against the codebook's inclusion and exclusion criteria; reported theme frequencies reflect the consensus coding. The full per-theme coefficients and the consensus coding matrix are provided in Online Appendix C.

Results

Descriptive statistics and overall effects

Substantial pre-post changes in student awareness emerged across all measured dimensions. In Cohort 2024–2 ($n = 43$), pretest scores (Table 3) revealed moderate baseline awareness in Valorization (Dimension 1: $M = 29.60$, $SD = 5.72$, representing 74% of

Table 3. Descriptive statistics – cohort 2024–2

Variable	M_{pre}	M_{post}	SD_{pre}	SD_{post}	ΔM
D1: Originalidad/valorization	29.60	31.74	5.72	5.53	2.14
D2: Self-awareness	24.77	29.14	6.37	6.51	4.37
D3: Applied comprehension	21.81	27.93	5.67	6.27	6.12
Total score	76.19	88.81	10.85	11.78	12.63

Note(s): Cohort 2024–2 ($n = 43$). Each dimension was measured using 8 items on a 5-point Likert scale (1 = strongly disagree to 5 = strongly agree); subscale range = 8–40, total scale range = 24–120. $\Delta M = M_{post} - M_{pre}$
Source(s): Authors' own work

maximum), Self-Awareness (Dimension 2: $M = 24.77$, SD = 6.37, 62%) and Applied Comprehension (Dimension 3: $M = 21.81$, SD = 5.67, 55%).

Post-intervention, all dimensions increased significantly: D1 to $M = 31.74$ (SD = 5.53), D2 to $M = 29.14$ (SD = 6.51) and D3 to $M = 27.93$ (SD = 6.27), yielding a total score improvement from $M = 76.19$ (SD = 10.85) to $M = 88.81$ (SD = 11.78).

Cohort 2025-1 (n = 48) exhibited a comparable baseline profile and parallel post-intervention improvements (Table 4).

The convergence of baseline profiles and change magnitudes across independent cohorts strengthens the robustness of the observed effects.

Inferential analysis: significance and effect sizes

Wilcoxon signed-rank tests (Table 5) confirmed statistically significant improvements ($p < 0.001$) across all dimensions and total scores in both cohorts, with exceptionally large effect sizes ($d \approx 2.81$–3.85 across dimensions; full statistics in Table 5).

These effect sizes substantially exceed conventional benchmarks for educational interventions ($d = 0.40$–0.60; Hattie (2009)). Their consistency across independent cohorts reduces the likelihood of sampling artifacts or cohort-specific confounds. Interpretation nonetheless requires recognizing that these metrics reflect changes i*n awareness and perception* of competencies rather than observed behavioral competence.

Dimensional pattern: differential impact across competency awareness types

A consistent hierarchical pattern emerged across both cohorts (Figure 2): Applied Comprehension showed the largest mean changes ($\Delta M \approx 6.1$–6.9), followed by Self-Awareness ($\Delta M \approx 4.4$–4.6)

Table 4. Descriptive statistics – cohort 2025–1

Variable	M_{pre}	M_{post}	SD_{pre}	SD_{post}	ΔM
D1: Originalidad/valorization	29.00	31.08	5.62	5.61	2.08
D2: Self-awareness	24.23	28.83	5.75	6.33	4.60
D3: Applied comprehension	20.15	27.06	5.07	4.94	6.92
Total score	73.38	86.98	9.37	11.45	13.60

Note(s): Cohort 2025–1 ($n = 48$). Each dimension was measured using 8 items on a 5-point Likert scale (1 = strongly disagree to 5 = strongly agree); subscale range = 8–40, total scale range = 24–120. $\Delta M = M_{post} - M_{pre}$
Source(s): Authors' own work

Table 5. Pre-post analysis by dimension and cohort (Wilcoxon signed-rank test)

Cohort	Variable	M_{pre}	M_{post}	ΔM	SD_{pre}	SD_{post}	p	Effect size
2024–2	D1: Originalidad/valorization	29.60	31.74	2.14	5.72	5.53	< 0.001	$r = 0.845$ ($d \approx 3.16$)***
	D2: Self-awareness	24.77	29.14	4.37	6.37	6.51	< 0.001	$r = 0.849$ ($d \approx 3.22$)***
	D3: Applied comprehension	21.81	27.93	6.12	5.67	6.27	< 0.001	$r = 0.877$ ($d \approx 3.66$)***
	Total score	76.19	88.81	12.63	10.85	11.78	< 0.001	$r = 0.874$ ($d \approx 3.60$)***
2025–1	D1: Originalidad/valorization	29.00	31.08	2.08	5.62	5.61	< 0.001	$r = 0.815$ ($d \approx 2.81$)***
	D2: Self-awareness	24.23	28.83	4.60	5.75	6.33	< 0.001	$r = 0.874$ ($d \approx 3.60$)***
	D3: Applied comprehension	20.15	27.06	6.92	5.07	4.94	< 0.001	$r = 0.887$ ($d \approx 3.85$)***
	Total score	73.38	86.98	13.60	9.37	11.45	< 0.001	$r = 0.874$ ($d \approx 3.59$)***

Note(s): Cohort 2024–2 ($n = 43$); Cohort 2025–1 ($n = 48$). $\Delta M = M_{post} - M_{pre}$. Effect sizes are reported as the Wilcoxon matched-pairs rank-biserial correlation r (Rosenthal, 1991), with approximate conversion to Cohen's d using $d \approx 2r/\sqrt{(1 - r^2)}$ (Borenstein et al., 2009) for comparability with educational intervention benchmarks. Within-subjects conversion inflates *d* relative to between-subjects equivalents (Lakens, 2013); the converted values are therefore reported for interpretive benchmarking and should not be compared directly against between-groups effect sizes. *** $p < 0.001$
Source(s): Authors' own work

and Valorization ($\Delta M \approx 2.1$). This D3 > D2 > D1 ordering replicates precisely across independent samples, indicating systematic differential impact rather than random variation.

The attenuated change in valorization reflects a ceiling effect: students entered with already favorable attitudes (72% of maximum), likely attributable to prior educational socialization or social desirability. The improvement in self-awareness suggests the exercise revealed discrepancies between students' assumed self-knowledge and their actual performance under pressure – a Dunning-Kruger recalibration from unconscious toward conscious incompetence. The maximum effect on Applied Comprehension (d > 3.6) indicates that the largest awareness shift occurred in the when and how of competency deployment, the dimension most directly addressed by the four-phase protocol. Students transitioned from declarative knowledge to procedural awareness under realistic constraints.

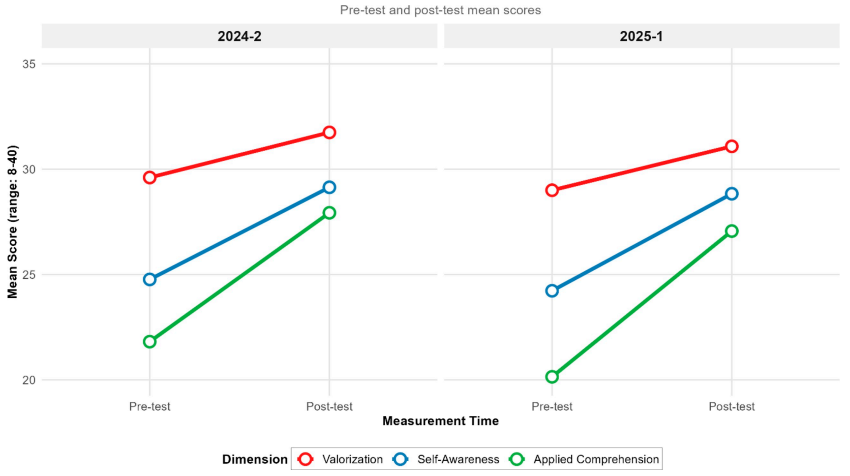

Figure 2. Evolution of competency dimension across time and cohort

Table 6. Pressure management shows maximum effects across all dimensions

Dimension	Subfactor	2024–2 (d)	2025–1 (d)	Pattern
D1: Originalidad/valorization	Leadership	1.60***	1.42***	Smallest effects
	Communication	1.64***	1.62***	(ceiling: Mpre ≈ 7.3)
	Group decisions	1.79***	1.34***	
	Pressure tolerance	2.34****	1.50****	+31–75% versus others
D2: Self-awareness	Leadership	2.60***	2.92***	Moderate effects
	Communication	1.93***	2.16***	
	Group decisions	2.97***	2.47***	
	Pressure tolerance	3.06****	3.05****	+3–58% versus others
D3: Applied comprehension	Leadership	3.30***	4.40***	Largest effects
	Communication	3.04***	4.31***	(lowest baseline)
	Decisions	3.56***	3.99***	
	Pressure management	4.16****	4.56****	+15–37% versus others

Note(s): All $p < 0.001$. Bold rows indicate pressure subfactors. Within-dimension comparison shows pressure competencies consistently dominate. Cross-cohort variance < 10% (SF3.4:4.16 versus 4.56 = 9.6% difference)
Source(s): Authors' own work

Subfactor analysis: temporal pressure as pedagogical mechanism

To examine candidate mechanisms underlying the dimensional pattern, we conducted subfactor-level analysis across all 24 items. Table 6 presents pressure-related subfactors alongside representative non-pressure subfactors from each dimension. This selective presentation highlights the theoretically critical pattern: pressure competencies systematically dominate within-dimension comparisons, a pattern consistent with temporal constraint, rather than generic engagement, as a candidate driver of the observed awareness shifts.

Figure 3 visualizes the complete subfactor hierarchy, revealing a systematic escalation from Valorization through Self-Awareness to Applied Comprehension, where pressure subfactors reach the upper extreme (d up to 4.56 in SF3.4) within each dimension.

The 15-minute constraint forced confrontation between theoretical knowledge (accessible under unlimited time) and procedural competence (executable under cognitive load). Students defaulted to heuristics despite possessing correct frameworks, as qualitative data illustrate: "I put chocolate high because 'we need energy now,' ignoring it's only 2 bars for 7 people for weeks." The modest D1 effects ($d \approx 2.81$–3.16 at the dimension level; subfactor range $d \approx 1.34$–2.34) versus extreme D3 effects ($d \approx 3.66$–3.85 at the dimension level; subfactor range $d \approx 3.04$–4.56) are consistent with the exercise functioning as revealer rather than developer: students did not alter values (already endorsed abstractly) but gained acute awareness of execution complexity under realistic constraints.

Cohort replication and robustness

The near-identical replication of results across Cohorts 2024 - 2 and 2025–1 (Figure 4) provides critical evidence of robustness effect.

Dimensional ordering (D3 > D2 > D1), the subfactor hierarchy and effect magnitudes converged despite independent sampling, different semester contexts and inevitable variation in group dynamics, making single-cohort artifacts, instructor effects or temporal confounds unlikely as primary explanations. This replication supports the protocol as reliably producing awareness shifts when implemented with fidelity, though generalization remains bounded: both cohorts derived from the same institution, academic level and cultural context. Replication across institutions, populations and cultural settings would establish broader external validity.

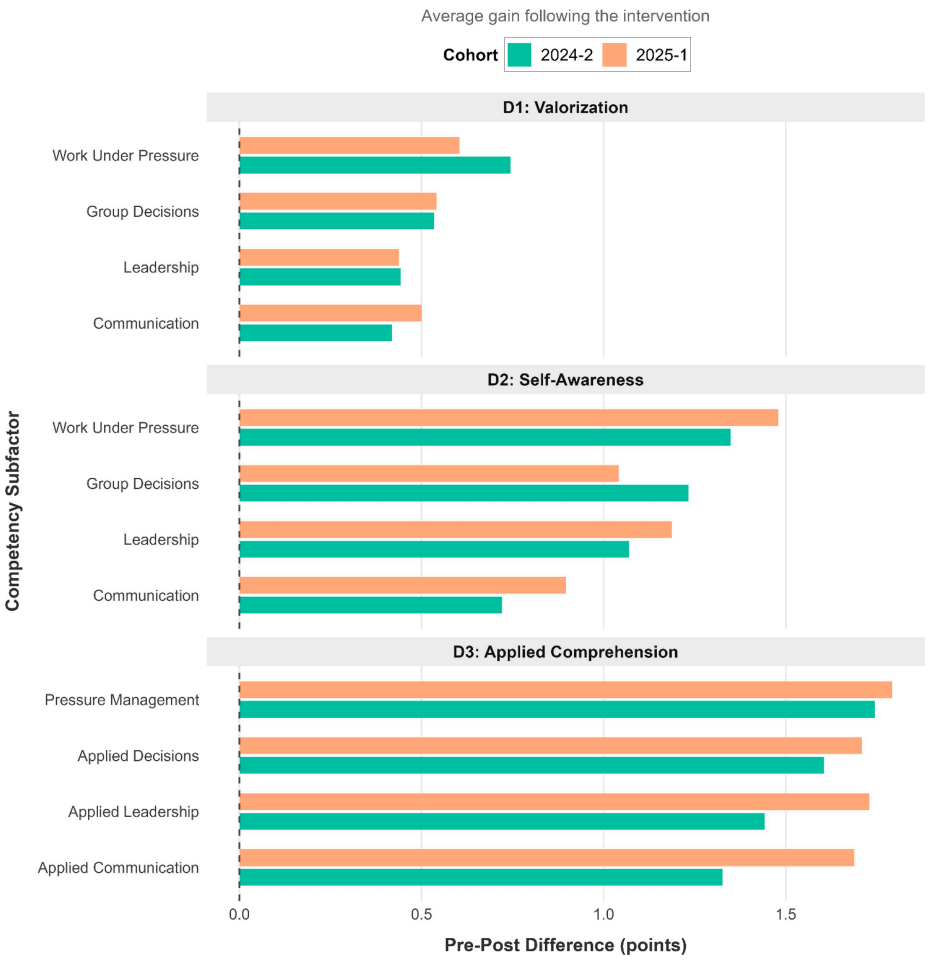

Figure 3. Mean pre-post differences by competency subfactor

Notes: Three patterns are consistent with temporal pressure as a candidate mechanism. First, within-dimension pressure dominance: pressure subfactors systematically exceed non-pressure items within identical dimensions, a finding at odds with a generic engagement account, which would predict uniform improvement. Second, baseline-effect inversion: the lowest baseline awareness produced the maximum gains whereas the highest baselines showed minimal gains, a pattern consistent with revelation rather than development. Third, cross-cohort replication: the pressure-dominance hierarchy replicated with under 10% variance, rendering sampling artifacts or instructor effects unlikely

Source: Authors' own work

Exercise performance analysis: decision-making patterns and strategic choices

Beyond self-reported awareness, actual performance data provide concrete evidence of challenges that generated documented awareness shifts. Performance patterns validate the knowing-doing gap students subsequently recognized in post-intervention measures.

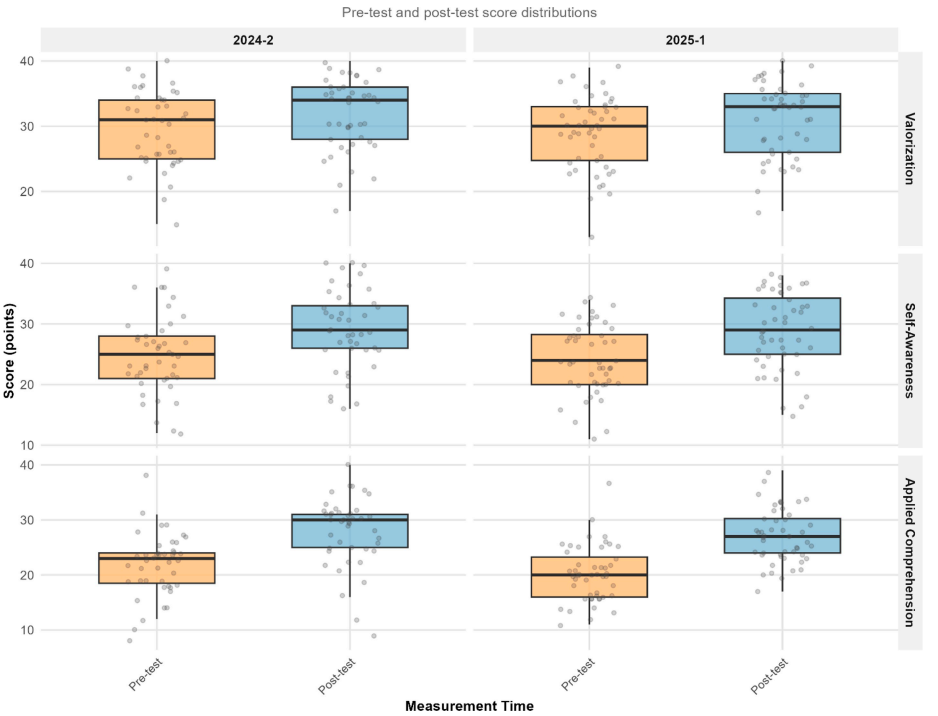

Figure 4. Distribution of scores by dimension and cohort

Source: Authors' own work

Cross-cohort performance patterns

Team synergy ranged from −8.17 to + 18.5, demonstrating that competent individuals do not guarantee superior collective decisions. Cohort 2024–2 showed limited synergy, with only one team achieving substantial gains; Cohort 2025–1 improved, with five teams generating positive synergy, the strongest transforming weak individual performance into competitive group decisions.

Survival rates doubled across cohorts (17% → 38%), yet 62% still failed the survivability threshold (≤60 points), a pattern suggesting systematic overestimation of crisis judgment capacity. Strategic choice patterns revealed action bias: 8 of 12 teams chose "rowing" despite rational analysis favoring "waiting" to conserve energy and remain locatable for rescue. This preference for active solutions over strategic patience is consistent with difficulty in suspending urgency in favor of probabilistic reasoning, the cognition-action tension students later articulated in qualitative interviews.

Leadership style correlated with performance: autocratic approaches clustered among worst individual scores (Teams 1, 3, 4, 6 in 2024–2), while democratic processes enabled knowledge integration (Team 6 in 2025–1: best group score 48 despite lacking strongest individuals). However, democratic style alone proved insufficient; several democratic teams failed to capitalize on perspective diversity, suggesting process structure matters beyond style.

Three tensions emerged: (1) Leadership-outcome dissociation, style does not guarantee integration effectiveness; (2) Cognition-action bias, preference for activity over strategy despite contrary evidence; (3) Individual-collective gap: aggregation of individual competence

does not predict collective decision quality. These patterns validate the self-reported awareness gains: students confronted genuine competency limitations during the exercise (median survival rate 28%), creating epistemic conditions for metacognitive revelation during debriefing. The improvement between cohorts (17% → 38% survival) suggests learning potential exists, but persistent gaps (62% failure rate) confirm that awareness-building requires structured reflection beyond mere exposure.

Qualitative evidence: student testimonies illuminating quantitative changes

Post-intervention semi-structured interviews ($n = 24$, purposively sampled across performance quartiles and synergy bands) illuminated the experiential mechanisms underlying the quantitative awareness gains. Independent coding by two researchers yielded substantial agreement ($\kappa = 0.73$; see Online Appendix C) and the six themes that emerged map systematically onto the quantitative dimensions and subfactors rather than standing apart from them: applied comprehension ($n = 18$) onto the D3 effects, self-awareness ($n = 21$) onto D2, pressure management ($n = 22$) onto the SF3.4 subfactor and collaborative modes ($n = 14$) onto the team-synergy variable (full mapping in Appendix C). The accounts therefore do not merely accompany the statistical results; they specify the mechanisms through which each measured shift occurred. We present them organized by theme, linking each to its quantitative referent. Table 7 provides a global picture of the thematic structure.

The six themes are presented in detail below, organized in descending order of conceptual centrality to the awareness construct, each illustrated by representative student accounts.

Applied comprehension: from intuition to systematic reasoning

The substantial improvement in Applied Comprehension ($d \approx 3.66–3.85$) manifested in students' recognition of gaps between intuitive assertions and evidence-based justification:

> In my individual ranking, I put the water tank first because obviously you need water to survive. But when we discussed as a team, someone said 'yes, but the instructions say we're 1000 miles

Table 7. Overview of qualitative themes and their quantitative referents

Theme	Participants (n)	Quantitative referent	Mechanism illuminated
Applied comprehension	18 / 24	D3 dimension ($d \approx 3.66–3.85$)	Recognition of the knowing–doing gap
Self-awareness	21 / 24	D2 dimension ($d \approx 3.22–3.60$)	Identification of unexamined behavioral patterns
Process consciousness	16 / 24	D2/D3 (process facet)	Dual-level monitoring of task and group dynamics
Pressure management	22 / 24	SF3.4 subfactor ($d \approx 4.16–4.56$)	Cognitive degradation under temporal constraint
Collaborative modes	14 / 24	Team-synergy variable	Interpersonal behaviors driving synergy outcomes
Affective regulation	12 / 24	SF2.4 (affective facet)	Emotional response under time pressure

Note(s): Themes derived through reflexive thematic analysis with inter-coder reliability check (pooled Cohen's $\kappa = 0.73$; 88.2% agreement). Detailed coefficients, the consensus matrix and the codebook are reported in Online Appendix C
Source(s): Authors' own work

> from shore, we'll die of dehydration long before rescue even if we ration water. The mirror can signal passing ships in one day.' I realized I was just saying 'it's logical' without thinking through the scenario constraints. Now I understand why managers need data, not just gut feeling." (Student 19, Team 2, individual score 58, team score 75, synergy-11).

Student 19's narrative reveals awareness of anchoring to survival heuristics without contextual analysis and conflating subjective confidence with objective reasoning. The exercise rendered visible the distinction between declarative knowledge ("water is important") and procedural competence ("how to systematically evaluate resource utility given constraints"). Despite negative team synergy, the student recognized the cognitive mechanism underlying superior collective decisions.

Self-awareness: recognizing behavioral patterns

Self-Awareness improvement ($d \approx 3.22–3.60$) reflected recognition of previously invisible coordination failures:

> We were all talking over each other about whether to row or stay. Someone yelled 'we should DO something!' and everyone agreed to row. But afterward, the instructor showed that rowing wastes energy, makes us harder to find, and we'd never reach shore 1000 miles away. I realized our group has an action bias, we chose activity over strategy because waiting felt passive. In administration class we learned about this bias, but I didn't know I actually had it until I saw us do it." (Student 31, Team 3, team score 60).

Student 31's team ultimately chose "wait" (the rational strategy) yet experienced action bias pressure during deliberation. Recognition occurred through experiencing the temptation to choose suboptimally, then consciously overriding it. The 15-minute constraint amplified this bias, making visible a cognitive pattern dormant in low-pressure contexts.

Temporal pressure as revelatory mechanism

Consistent with maximum effects in pressure management ($d = 4.16–4.56$), students identified time constraints as exposing hidden incompetencies:

> With unlimited time, I could analyze each object carefully. But with 15 minutes, I panicked and just picked based on what seemed urgent. I put the chocolate high because 'we need energy now,' ignoring that it's only 2 bars for maybe 7 people for weeks. The exercise showed me I don't actually know how to prioritize when stressed. That's scary for real emergencies." (Student 08, Team 3, individual score 80, team score 76, synergy-7.6).

Time pressure induced shifts from analytical reasoning to heuristic-based judgment. The student recognizes this as a competency gap, not merely performance failure.

Collaborative dynamics: failure and success modes

Students in autocratic-led teams articulated awareness of how individual behaviors undermined collective performance:

> I had the best individual score in my group [56], so I thought I should lead. I kept saying 'trust me, put the water tank at #1, I'm sure.' But we got 62 points as a team, worse than my individual 56. The instructor said I was being autocratic and not listening to others' ideas. One teammate had put the mirror high with good reasoning about signaling, but I dismissed it. I learned that being confident doesn't mean being right, and that my confidence can actually make the team worse." (Student 04, Team 1, individual 56, team 62, synergy-2.43).

Conversely, high-synergy teams recognized productive mechanisms:

> Our team made a rule: everyone gets 1 minute uninterrupted to explain their top 3 items with reasons. Then we voted. When it was [Student 48]'s turn, she said 'the nylon sheet because it can collect rainwater, make shelter, and signal with reflection.' Nobody had thought about multi-use items. We ended up scoring 48 as a team when our individual average was 61, we gained 13 points just from listening to everyone fully." (Student 44, Team 6, individual 52, team 48, synergy +13.4).

Team 6 had the best performance despite lacking the strongest individual performers, revealing how structured turn-taking enabled knowledge integration and systematic thinking.

Debriefing as catalytic mechanism

Students credited post-exercise debriefing as the moment of insight formation:

> During the exercise, I was just trying to survive, rank the items, convince my team, beat the clock. I didn't think about how I was doing it. But when the professor asked 'Who dominated the conversation? Who stayed quiet? How did you handle disagreements?', that's when I realized I had interrupted people three times. I didn't notice during the exercise, but the debriefing made it visible. The learning happened afterward, not during." (Student 39, Team 1, individual 71, team 60, synergy-0.5).

Exercise creates raw experience; debriefing transforms it into learning through guided reflection. Student 39's claim that "learning happened afterward, not during" demonstrates experiential exercises require explicit metacognitive scaffolding.

These accounts illuminate how statistical improvements ($d > 2.8$) occurred through: discrepancy exposure between self-perception and observable behavior, contextualized application connecting abstract principles to concrete scenarios, process visualization rendering invisible dynamics consequential and metacognitive scaffolding redirecting attention from outcomes to processes.

The testimonies validate the quantitative pattern (D3 > D2 > D1): largest changes in applied comprehension revealed the knowing-doing gap; moderate changes in self-awareness as students recognized behavioral patterns; smallest changes in valorization due to ceiling effects. These data support interpretation of effect sizes as awareness revelation rather than competency development. The shipwreck exercise functions as a diagnostic intervention, creating epistemological humility necessary for subsequent formal instruction.

Discussion

This study examined whether and how a brief survival simulation shifts students' awareness of managerial competencies. We organize the discussion around the three research questions:

RQ4. The magnitude and nature of awareness change.

The large pre-post effects ($d = 2.81–3.85$ at the dimension level) indicate that the Shipwreck exercise functions as a consciousness-raising intervention rather than a skill-building program. The magnitudes observed are unusually large for educational interventions and warrant cautious interpretation, which we develop below. Within 90 min students did not acquire managerial competencies; they became aware of the complexity of applying theoretical knowledge under realistic constraints, of discrepancies between assumed and actual capability and of the role of time pressure in decision quality. This interpretation reconciles the unusually large effects with pedagogical theory: awareness can shift rapidly through a single disruptive experience (Mezirow, 1991), whereas behavioral competence requires sustained practice. The exercise induced what Schön (1983) termed reflection-in-action, students executed decisions

while observing their own patterns under stress, producing immediate metacognitive insight. Three considerations indicate that the magnitude reflects theoretically expected processes rather than artifact. First, the construct measured is awareness revelation (a qualitative state change), not skill accumulation; transitions from unconscious to conscious incompetence can occur through singular disruptive experiences (Kitchenham, 2008; Taylor, 2007). Second, the within-subjects design inflates standardized differences relative to between-subjects equivalents, as noted in the analysis (Lakens, 2013); both indices are reported for transparency. Third, part of the observed shift likely reflects internal-standard recalibration (response shift; Howard (1980); Sprangers and Schwartz (1999)) students initially evaluated "working under pressure" against academic deadlines, then against survival decision-making. Critically, this recalibration is itself the intended learning, the acquisition of realistic benchmarks for managerial demand, rather than a confound to be explained away:

RQ5. The differential pattern and temporal pressure as candidate mechanism.

The dimensional hierarchy (D3 > D2 > D1), replicated across cohorts, is consistent with temporal pressure as a candidate mechanism rather than with the gamified frame in general. Applied comprehension gained most precisely where baseline awareness was lowest and where execution under load most exposed the knowing–doing gap; valorization gained least, constrained by ceiling effects on already-endorsed beliefs. The dominance of the pressure-management subfactor (SF3.4, $d = 4.16$–4.56) points to the 15-minute constraint as a likely *locus* of the effect: under cognitive load learners default to heuristics and their default patterns become observable to themselves (Kahneman, 2011), generating the disorienting dilemma that activates transformative learning (Mezirow, 1991). Generic engagement account would predict uniform improvement across dimensions; the observed specificity does not:

RQ6. Experiential processes underlying the change.

The qualitative accounts converge with the quantitative pattern and specify its mechanism. Students described time-pressure-induced exposure of heuristic reasoning, recognition of previously invisible coordination failures and the gap between abstract endorsement and situated execution, processes consistent with a diagnostic rather than skill-developing function. The convergence of self-reported mechanism with the dimensional and subfactor hierarchies strengthens the interpretation that the intervention reveals competency complexity rather than building competence within the session.

Contribution to international business education

The study contributes to the field of international business education in three connected ways. *Theoretically*, it distinguishes competency *awareness* from competency *development* as separable pedagogical outcomes and shows that a brief, intensive simulation can produce the former without claiming the latter. This distinction refines how experiential methods are understood in management curricula: rather than asking whether a 90-minute exercise "develops skills", a claim the evidence would not support, it reframes the exercise as a diagnostic that establishes the conscious incompetence (Gordon and Burch, 1974) on which subsequent skill-building depends. The three-dimensional operationalization (valorization, self-awareness, applied comprehension) offers other educators a transferable way to measure this outcome rather than inferring it.

Methodologically, the study demonstrates how a two-cohort replication can supply analytical robustness when an active control group is neither pedagogically nor ethically available, a constraint common to authentic classroom research in business programs. By

holding the protocol constant and examining whether dimensional and subfactor hierarchies recur under independent sampling, the design converts a limitation of practitioner research into a source of inferential strength, offering a template for instructors who study their own teaching within the action-research tradition (Coghlan, 2010; Starr-Glass, 2011).

Contextually, the study extends the evidence base for experiential management pedagogy to an emerging-market setting where such methods remain pedagogically novel. Most evidence on simulations derives from North American and European business schools where experiential learning is routine; the present findings suggest that in contexts dominated by lecture-based instruction, the disruptive contrast of a first experiential encounter may be precisely what makes awareness salient. This positions the contribution not as a claim of universal efficacy but as a boundary-conditioned insight relevant to comparable international contexts, Latin American, South Asian, Sub-Saharan African and Middle Eastern business programs, where the same pedagogical novelty obtains. For international business education specifically, the competencies the exercise renders visible (decision-making under uncertainty, coordination under pressure, recognition of one's own default heuristics) are those that cross-cultural and globally distributed managerial work most demands.

Pedagogical implications

The differential dimensional effects (D3 > D2 > D1), the dominance of pressure-management subfactors and the qualitative evidence on debriefing as the *locus* of learning translate into four concrete design recommendations for instructors integrating the Shipwreck exercise into management curricula.

First, position the exercise as a transition intervention. The findings indicate that maximum impact occurs when students already hold foundational theoretical knowledge but lack experiential grounding. Optimal placement is therefore after the introductory module on managerial competencies (typically weeks 3–5 of a first-year course) and before applied skill-building modules on decision-making, leadership or team coordination. Using the exercise as an opening session for the term forfeits the diagnostic contrast that produces awareness; using it after applied modules forfeits its preparatory function.

Second, preserve the 15-minute time constraint with fidelity. The pressure-management subfactors carried the strongest effects ($d \approx 4.16$–4.56), a pattern consistent with temporal pressure as the candidate active mechanism. Extending the individual ranking phase beyond 15 min or allowing untimed group deliberation, would likely attenuate the awareness yield by removing the cognitive load that exposes students' default heuristics. Instructors who must adapt the timing to a 90-minute session should compress phases 3 and 4 rather than relax the constraint in phase 1.

Third, structure the debriefing around process rather than outcome. The qualitative accounts consistently identified debriefing – not the ranking task itself – as the moment when awareness crystallized. Rather than asking "which team got the closest to the expert ranking?" instructors should anchor the debriefing in four metacognitive prompts derived from the themes we identified: How did you allocate the 15 min? Where did your reasoning shift from analysis to heuristics? Whose contributions were heard and whose were not? Which assumption about managerial work did this experience disconfirm? A 30–40 min debriefing using these prompts is more pedagogically productive than a longer one centered on the "correct" ranking.

Fourth, link the exercise explicitly to subsequent modules. Awareness without follow-through decays. Instructors should reference specific Shipwreck moments when introducing later content on coordination, conflict management, time-pressure decision-making and leadership style. This sequencing converts the diagnostic into a recurring reference point, deepening the awareness gains into durable conceptual scaffolding across the term.

Limitations and boundary conditions

Four limitations bound these findings. First, the absence of an active control condition limits causal inference: we cannot fully exclude generic experiential learning, Hawthorne or regression effects. Three patterns nonetheless render such explanations less parsimonious. The dimensional hierarchy (D3 > D2 > D1) is theoretically coherent with the pressure-revelation mechanism, whereas generic engagement would predict uniform improvement; the pressure-management subfactor dominates ($d = 4.16–4.56$), implicating the temporal constraint rather than the gamified frame; and qualitative accounts reference mechanism-specific exposure ("I panicked and just picked based on what seemed urgent") rather than generic enjoyment. These do not substitute for experimental control, which future randomized designs with active comparison conditions (e.g. the identical scenario with extended, unpressured deliberation) should supply.

Second, reliance on self-report exposes the findings to demand characteristics and social desirability and to *response shift*, the recalibration of internal standards that can inflate pre-post differences (Howard, 1980; Sprangers and Schwartz, 1999). Two considerations partially mitigate this. The direction of several findings runs against social desirability: students reported greater recognition of their own gaps, not greater competence. And, as argued under *RQ1*, standard recalibration is partly the intended learning rather than pure bias. Still, convergent validation against behavioral indicators (peer ratings, observer coding of subsequent simulations) remains necessary and a retrospective then-test design would help separate response shift from absolute change.

Third, single-institution sampling bounds generalizability. The setting, a private Ecuadorian university where experiential methods are pedagogically novel, may amplify effects through disruptive contrast with lecture-based norms. We treat this as an intentional boundary condition rather than a flaw: the study establishes proof of concept for high-novelty contexts, not universal efficacy. Effects may attenuate where simulation-based learning is routine and the threshold for a disorienting dilemma is higher and may differ where cultural norms attach greater stigma to public performance failure (Van Rompay-Bartels and Geessink, 2021). The two-cohort replication demonstrates robustness to sampling variation within this context, not generalization across institutional types; extension to public universities and other national systems requires direct empirical testing.

Fourth, the design measured immediate post-intervention awareness only. Whether these shifts persist, decay or translate into later behavioral competence is unknown. Beyond the controlled and longitudinal designs implied above, three further directions would advance the agenda: mediational testing of the hypothesized pathway (temporal pressure → self-observed performance gaps → awareness gain); behavioral triangulation linking awareness to downstream decision quality; and cross-cultural replication to identify the cultural moderators that bound effect magnitude across comparable international contexts.

Conclusions

This study asked whether and how a brief survival simulation alters first-year students' reported awareness of managerial competencies in an emerging-market business program. Across two independent cohorts, the Shipwreck exercise was associated with large, replicable pre-post gains in awareness, with the clear dimensional hierarchy of applied comprehension over self-awareness over valorization (*RQ4*). This pattern, together with the dominance of the pressure-management subfactor, is consistent with temporal pressure, rather than the gamified frame in general, as a candidate mechanism that renders the knowing–doing gap visible (*RQ5*) and the students' own accounts of heuristic exposure and coordination failure under time constraint converge with that interpretation (*RQ6*). Taken

together, the findings support a reframing of brief, intensive simulations: their pedagogical value lies not in developing managerial competence within the session but in revealing its complexity, establishing the conscious incompetence on which later skill-building depends. For international business education, this distinction matters most in contexts where experiential methods are still novel, where a first disorienting encounter can make competency gaps salient in ways routine lecture instruction does not. Positioned deliberately between foundational theory and applied skill modules and preserved with fidelity to its temporal constraint, the exercise offers educators across comparable international settings a low-cost, replicable diagnostic that prepares students for the competency development that follows. The contribution is therefore both conceptual, separating awareness from competence as outcomes and practical, offering a transferable design for emerging-market management classrooms.

Acknowledgements

A Large Language Model (Claude Opus 4.7, Anthropic) was used as an editorial aid during the writing and revision stages, specifically for English translation and language revision, and for editorial assistance in the preparation of the manuscript. The authors reviewed, edited and took full responsibility for all content of the final manuscript.

References

Ajzen, I. (1991), "The theory of planned behavior", *Organizational Behavior and Human Decision Processes*, Vol. 50 No. 2, pp. 179-211, doi: 10.1016/0749-5978(91)90020-T.

Almeida, F. and Simões, J. (2019), "The role of serious games, gamification and industry 4.0 tools in the education 4.0 paradigm", *Contemporary Educational Technology*, Vol. 10 No. 2, pp. 120-136, doi: 10.30935/cet.554469.

Armenia, S., Barnabè, F., Nonino, F. and Pompei, A. (2024), "Improving project management skills by integrating a boardgame into educational paths", *The International Journal of Management Education*, Vol. 22 No. 2, p. 100969, doi: 10.1016/j.ijme.2024.100969.

Ashforth, B.E., Harrison, S.H. and Corley, K.G. (2008), "Identification in organizations: an examination of four fundamental questions", *Journal of Management*, Vol. 34 No. 3, pp. 325-374, doi: 10.1177/0149206308316059.

Barbosa, M.W. and de Ávila Rodrigues, C. (2020), "Project portfolio management teaching: Contributions of a gamified approach", *The International Journal of Management Education*, Vol. 18 No. 2, p. 100388, doi: 10.1016/j.ijme.2020.100388.

Bedwell, W.L., Fiore, S.M. and Salas, E. (2014), "Developing the future workforce: an approach for integrating interpersonal skills into the MBA classroom", *Academy of Management Learning and Education*, Vol. 13 No. 2, pp. 171-186, doi: 10.5465/amle.2011.0138.

Borenstein, M., Hedges, L.V., Higgins, J.P.T. and Rothstein, H.R. (2009), "Converting among effect sizes", in Borenstein, M., Hedges, J. L.V., Higgins, P.T. and Rothstein H.R. (Eds), *Introduction to meta-Analysis*, John Wiley and Sons, pp. 45-49.

Boyatzis, R. (1982), *The Competent Manager: A Model for Effective Performance*, John Wiley and Sons.

Braun, V. and Clarke, V. (2006), "Using thematic analysis in psychology", *Qualitative Research in Psychology*, Vol. 3 No. 2, pp. 77-101, doi: 10.1191/1478088706qp063oa.

Braun, V. and Clarke, V. (2019), "Reflecting on reflexive thematic analysis", *Qualitative Research in Sport, Exercise and Health*, Vol. 11 No. 4, pp. 589-597, doi: 10.1080/2159676X.2019.1628806.

Castillo-Parra, B., Hidalgo-Cajo, B.G., Vásconez-Barrera, M. and Oleas-López, J. (2022), "Gamification in higher education: a review of the literature", *World Journal on Educational Technology: Current Issues*, Vol. 14 No. 3, pp. 797-816, doi: 10.18844/wjet.v14i3.7341.

Cicchetti, D.V. and Feinstein, A.R. (1990), "High agreement but low kappa: II. Resolving the paradoxes", *Journal of Clinical Epidemiology*, Vol. 43 No. 6, pp. 551-558, doi: 10.1016/0895-4356(90)90159-M.

Coghlan, D. (2010), "Seeking common ground in the diversity and diffusion of action research and collaborative management research action modalities: toward a general empirical method", in *Research in Organizational Change and Development*, Emerald Group Publishing Limited, Bingley, pp. 149-181, doi: 10.1108/S0897-3016(2010)0000018009.

Coghlan, D. and Jacobs, C. (2005), "Kurt Lewin on reeducation: Foundations for action research", *The Journal of Applied Behavioral Science*, Vol. 41 No. 4, pp. 444-457, doi: 10.1177/0021886305277275.

Crespo-Martínez, E., Bueno, S. and Gallego, M.D. (2025), "Uses and gratifications approach to evaluate the continuance intention of ATIC: a serious video game to learn entrepreneurship", *The International Journal of Management Education*, Vol. 23 No. 2, p. 101125, doi: 10.1016/j.ijme.2024.101125.

Creswell, J.W. and Plano Clark, V.L. (2018), *Designing and Conducting Mixed Methods Research*, 3rd ed. SAGE Publications, Inc, Thousand Oaks, CA.

Crookall, D. (2010), "Serious games, debriefing, and simulation/gaming as a discipline", *Simulation and Gaming*, Vol. 41 No. 6, pp. 898-920, doi: 10.1177/1046878110390784.

Davila, A. and Elvira, M.M. (2012), "Humanistic leadership: Lessons from Latin America", *Journal of World Business*, Vol. 47 No. 4, pp. 548-554, doi: 10.1016/j.jwb.2012.01.008.

Deterding, S., Dixon, D., Khaled, R. and Nacke, L. (2011), "From game design elements to gamefulness: defining 'gamification"', *Paper presented at the Proceedings of the 15th International Academic MindTrek Conference: envisioning Future Media Environments*, ACM, Tampere, Finland, doi: 10.1145/2181037.2181040.

Durrani, U.K., Al-Naymat, G., Ayoubi, R.M., Kamal, M.M. and Hussain, H. (2022), "Gamified flipped classroom versus traditional classroom learning: which approach is more efficient in business education?", *The International Journal of Management Education*, Vol. 20 No. 1, p. 100595, doi: 10.1016/j.ijme.2021.100595.

Edland, A. and Svenson, O. (1993), "Judgment and decision making under time pressure", in Svenson, O. and Maule, A. J. (Eds), *Time Pressure and Stress in Human Judgment and Decision Making*, Springer US, Boston, MA, pp. 27-40.

Eurich, T. (2018), "What Self-Awareness really is (and how to cultivate it)", Harvard Business Review, (January 4), available at: https://hbr.org/2018/01/what-self-awareness-really-is-and-how-to-cultivate-it

Faria, A.J. (2001), "The changing nature of business simulation/gaming research: a brief history", *Simulation and Gaming*, Vol. 32 No. 1, pp. 97-110, doi: 10.1177/104687810103200108.

Flavell, J.H. (1979), "Metacognition and cognitive monitoring: a new area of cognitive–developmental inquiry", *American Psychologist*, Vol. 34 No. 10, pp. 906-911, doi: 10.1037/0003-066X.34.10.906.

Fritz, C.O., Morris, P.E. and Richler, J.J. (2012), "Effect size estimates: current use, calculations, and interpretation", *Journal of Experimental Psychology: General*, Vol. 141 No. 1, pp. 2-18, doi: 10.1037/a0024338.

Gatti, L., Ulrich, M. and Seele, P. (2019), "Education for sustainable development through business simulation games: an exploratory study of sustainability gamification and its effects on students' learning outcomes", *Journal of Cleaner Production*, Vol. 207, pp. 667-678, doi: 10.1016/j.jclepro.2018.09.130.

Gordon, T. and Burch, N. (1974), *T.E.T., Teacher Effectiveness Training*, David McKay Books, New York, NY.

Grayot, J.D., Beck, L. and Heijmeskamp, T. (2024), "Dual process theory and the challenges of functional individuation", *Phenomenology and the Cognitive Sciences*, pp. 1-23, doi: 10.1007/s11097-024-10000-3.

Guetterman, T.C., Fetters, M.D. and Creswell, J.W. (2015), "Integrating quantitative and qualitative results in health science mixed methods research through joint displays", *The Annals of Family Medicine*, Vol. 13 No. 6, p. 554, doi: 10.1370/afm.1865.

Hamari, J., Koivisto, J. and Sarsa, H. (2014), "Does gamification work? – a literature review of empirical studies on gamification", *Paper presented at the 2014 47th Hawaii International Conference on System Sciences*, IEEE Computer Society Press, Waikoloa, HI.

Hattie, J.A.C. (2009), *Visible Learning. A Synthesis of over 800 Meta-Analyses Relating to Achievement*, 1st Ed., Routledge, John Hattie.

Ho, S.C. and Chen, J.L. (2023), "Developing the e-commerce competency for entrepreneurship education from a gamified competition", *The International Journal of Management Education*, Vol. 21 No. 1, p. 100737, doi: 10.1016/j.ijme.2022.100737.

Howard, G.S. (1980), "Response-Shift bias: a problem in evaluating interventions with pre/post Self-Reports", *Evaluation Review*, Vol. 4 No. 1, pp. 93-106, doi: 10.1177/0193841X8000400105.

Hyams-Ssekasi, D. and Taheri, F. (2022), "Re-assessing entrepreneurship education and gamification as a learning process", in *Technology and Entrepreneurship Education: Adopting Creative Digital Approaches to Learning and Teaching*, Springer International Publishing, Cham, Switzerland, pp. 3-23.

Jerônimo Soares, A.M.J., Neves Bezerra de Melo, F.L.N.B., de Tarso Alves Dantas, S., da Silva, M.P. and Genuino, S.L.V.P. (2024), "Gamification in entrepreneurship education: a systematic literature review and future research agenda", *REGEPE Entrepreneurship and Small Business Journal*, Vol. 13 No. 2, p. e2389, doi: 10.14211/regepe.esbj.e2389.

Johnson, D.W. and Johnson, R.T. (1999), *Learning Together and Alone. Cooperative, Competitive and Individualistic Learning*, 5th Ed., Allyn and Bacon, Boston, MA.

Kahneman, D. (2011). *Thinking, Fast and Slow: Farrar*, Straus and Giroux, New York, NY.

Kapp, K.M. (2012), *The Gamification of Learning and Instruction: Game-Based Methods and Strategies for Training and Education*, Pfeiffer and Company, San Francisco, CA.

Kitchenham, A. (2008), "The evolution of john Mezirow's transformative learning theory", *Journal of Transformative Education*, Vol. 6 No. 2, pp. 104-123, doi: 10.1177/1541344608322678.

Kolb, D.A. (2015), *Experiential Learning: Experience As The Source Of Learning And Development*, 2nd Ed., Pearson Education, Upper Saddle, NJ.

Kothiyal, N., Bell, E. and Clarke, C. (2018), "Moving beyond mimicry: developing hybrid spaces in Indian business schools", *Academy of Management Learning and Education*, Vol. 17 No. 2, pp. 137-154, available at: www.jstor.org/stable/26528570

Lakens, D. (2013), "Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs", *Frontiers in Psychology*, Vol. 4, p. 863, doi: 10.3389/fpsyg.2013.00863.

Landis, J.R. and Koch, G.G. (1977), "The measurement of observer agreement for categorical data", *Biometrics*, Vol. 33 No. 1, pp. 159-174, doi: 10.2307/2529310.

Legaki, N.Z., Karpouzis, K., Assimakopoulos, V. and Hamari, J. (2021), "Gamification to avoid cognitive biases: an experiment of gamifying a forecasting course", *Technological Forecasting and Social Change*, Vol. 167, p. 120725, doi: 10.1016/j.techfore.2021.120725.

Liu, S., Lv, M., Huang, Q. and Wang, Y. (2024), "Professional, scholar, or knowledge worker? Identity construction of Chinese management researchers amid the research–practice gap", *Plos One*, Vol. 19 No. 8, p. e0306833, doi: 10.1371/journal.pone.0306833.

López, F.R., Arias-Oliva, M., Pelegrín-Borondo, J. and Marín-Vinuesa, L.M. (2021), "Serious games in management education: an acceptance analysis", *The International Journal of Management Education*, Vol. 19 No. 3, p. 100517, doi: 10.1016/j.ijme.2021.100517.

McGonigal, J. (2011), *Reality is Broken: Why Games Make us Better and How They Can Change the World*, Penguin Press, New York, NY.

Mason, K., Anderson, L., Black, K. and Roberts, A. (2024), "A shout-out for the value of management education research: 'pedagogy is not a dirty word", *British Journal of Management*, Vol. 35 No. 2, pp. 539-549, doi: 10.1111/1467-8551.12805.

Maule, A.J., Hockey, G.R.J. and Bdzola, L. (2000), "Effects of time-pressure on decision-making under uncertainty: changes in affective state and information processing strategy", *Acta Psychologica*, Vol. 104 No. 3, pp. 283-301, doi: 10.1016/S0001-6918(00)00033-0.

Memar, N., Sundström, A. and Larsson, T. (2021), "Teaching causation and effectuation in the large classroom: a production–trade game", *Journal of Management Education*, Vol. 45 No. 3, pp. 438-478, doi: 10.1177/1052562920951971.

Mezirow, J. (1991), *Transformative Dimensions of Adult Learning*, Jossey-Bass, San Francisco.

Mintzberg, H. (2004), *Managers, Not MBAs: A Hard Look at the Soft Practice of Managing and Management Development*, Berrett-Koehler Publishers, San Francisco.

Montero-Benavides, P., Albort-Morant, G. and Masero-Moreno, I.C. (2025), "Evaluating a gamified assessment model to enhance performance and engagement in finance education: a quasi-experimental study", *The International Journal of Management Education*, Vol. 23 No. 3, p. 101266, doi: 10.1016/j.ijme.2025.101266.

Nemiroff, P.M. and Pasmore, W.A. (1975), "Lost at sea: a consensus-seeking task", in Pfeiffer, J.W. and Jones, J.E. (Eds), *The 1975 Annual Handbook for Group Facilitators*, University Associates, La Jolla, CA, pp. 28-34.

Nemiroff, P.M., Pasmore, W.A. and Ford, D.L. (1976), "The effects of two normative structural interventions on established and ad hoc groups: Implications for improving decision making effectiveness", *Decision Sciences*, Vol. 7 No. 4, pp. 841-855, doi: 10.1111/j.1540-5915.1976.tb00723.x.

Nunnally, J. and Bernstein, I. (1994), *Psychometric Theory*, 3rd Ed., McGraw-Hill, New York, NY.

Pettigrew, A. and Starkey, K.E.N. (2016), "From the guest editors: the legitimacy and impact of business schools – key issues and a research agenda", *Academy of Management Learning and Education*, Vol. 15 No. 4, pp. 649-664, available at: www.jstor.org/stable/26400178

Rapoport, A. (1966), *Two-Person Game Theory; the Essential Ideas*, University of MI Press, Ann Arbor.

Rosenthal, R. (1991), *Meta-Analytic Procedures for Social Research*, SAGE Publications, Inc, Newbury Park, CA, doi: 10.4135/9781412984997.

Sailer, M., Hense, J.U., Mayr, S.K. and Mandl, H. (2017), "How gamification motivates: an experimental study of the effects of specific game design elements on psychological need satisfaction", *Computers in Human Behavior*, Vol. 69, pp. 371-380, doi: 10.1016/j.chb.2016.12.033.

Salen, K. and Zimmerman, E. (2003), *Rules of Play: Game Design Fundamentals*, The MIT Press, Cambridge, MA.

Schön, D. A. (1983). *The Reflective Practitioner: How Professionals Think in Action*, Basic Books, New York, NY.

Schraw, G. and Moshman, D. (1995), "Metacognitive theories", *Educational Psychology Review*, Vol. 7 No. 4, pp. 351-371, doi: 10.1007/BF02212307.

Siala, H., Tajvidi, M., Wang, Y., Hajli, N., Richard, M.O. and Brannan, M. (2025), "Using serious games for learning business ethics: Exploring its socio-technical enablers and cross-cultural influences", *Journal of Management Education*, Vol. 49 No. 4, pp. 570-607, doi: 10.1177/10525629251316539.

Silva, R.J.R., Gouveia-Rodrigues, R.G. and Leal, C.T.P. (2019), "Gamification in management education: a systematic literature review", *BAR - Brazilian Administration Review*, Vol. 16 No. 2, p. e180103, doi: 10.1590/1807-7692bar2019180103.

Sprangers, M.A.G. and Schwartz, C.E. (1999), "Integrating response shift into health-related quality of life research: a theoretical model", *Social Science and Medicine*, Vol. 48 No. 11, pp. 1507-1515, doi: 10.1016/S0277-9536(99)00045-3.

Starr-Glass, D. (2011), "Between stereotype and authenticity: using action research in a cross-cultural management course", *Journal of International Education in Business*, Vol. 4 No. 2, pp. 112-124, doi: 10.1108/18363261111189513.

Sutton, A., Williams, H.M. and Allinson, C.W. (2015), "A longitudinal, mixed method evaluation of self-awareness training in the workplace", *European Journal of Training and Development*, Vol. 39 No. 7, pp. 610-627, doi: 10.1108/EJTD-04-2015-0031.

Tashakkori, A., Johnson, R.B. and Teddlie, C. (2020), "Foundations of mixed methods research", *Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*, 2nd Ed. SAGE Publications, Inc, Newbury Park, CA.

Taylor, E.W. (2007), "An update of transformative learning theory: a critical review of the empirical research (1999–2005)", *International Journal of Lifelong Education*, Vol. 26 No. 2, pp. 173-191, doi: 10.1080/02601370701219475.

Triantafyllou, S.A., Georgiadis, C. and Sapounidis, T. (2025), "Gamification in education and training: a literature review", *International Review of Education*, Vol. 71 No. 3, pp. 483-517, doi: 10.1007/s11159-024-10111-8.

Van Rompay-Bartels, I. and Geessink, J. (2021), "Exploring peer feedback on behavior in the international classroom: a case study on students' experiences and perceptions", *Journal of International Education in Business*, Vol. 16 No. 1, pp. 1-17, doi: 10.1108/JIEB-07-2020-0063.

Yalcin, T., Pehlivan, E. and Nistor, C. (2020), "The impact of gamification on learner engagement, enjoyment and performance: an abstract", in *Developments in Marketing Science: Proceedings of the Academy of Marketing Science, Springer Nature, Cham, Switzerland*, pp. 289-290.

Yin, R.K. (2018), *Case Study Research and Applications: Design and Methods*, 6th Ed., Sage, Thousand Oaks, CA.

Zichermann, G. and Cunningham, C. (2011), *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*, O'Reilly Media, Inc, Sebastopol, CA.

Supplementary material

The supplementary material for this article can be found online.

Corresponding author

Gelmar García-Vidal can be contacted at: gelmar.garcia@ute.edu.ec