

Interaction Quality and Feedback Uptake Link Small Class Teaching With Deep Learning Through Student Engagement

Yiran Zhou

YiranZhou1983@163.com

Stanford University

Research Article

Keywords: small-class teaching, deep learning, interaction quality, feedback uptake, student engagement, three-wave design

Posted Date: September 2nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10467685/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Small-class teaching is often promoted as a practical route to better undergraduate learning. Yet a smaller class does not automatically produce deeper learning. This study examines how small-class instructional support is associated with deep learning through interaction quality, feedback uptake, and student engagement. Drawing on dialogic pedagogy, feedback literacy, student engagement theory, and the deep approach to learning, the study develops and tests a sequential mediation model. A three-wave mixed-methods design was used. Survey data were collected from 612 undergraduate students nested in 38 classes and 27 courses across four Chinese universities. Classroom observations, assignment-revision coding, and semi-structured interviews with 28 students and 10 instructors were also collected. Confirmatory factor analysis supported the five-factor measurement model. Structural equation modeling showed that small-class instructional support was positively associated with interaction quality, which predicted feedback uptake; feedback uptake then predicted student engagement, and student engagement predicted deep learning. Bootstrap results supported the theoretically specified sequential indirect pathway. Multilevel SEM, group comparison, classroom observation, assignment-revision evidence, and interview themes produced convergent support for the mechanism. The findings suggest that small-class teaching works less as a structural condition and more as an opportunity structure. Its value depends on whether instructors use the smaller setting to build dialogue, make feedback usable, and sustain student engagement. The study contributes to higher education research by shifting attention from class size itself to the learning processes through which small classes become deep-learning environments.

1. Introduction

Small-class teaching has become a common reform strategy in undergraduate education. Universities use it to address weak interaction, passive attendance, limited feedback, and the distance between teachers and students in large lecture-based courses. The assumption is simple: when the number of students is reduced, teachers can know students better, students can speak more often, feedback can become more personal, and learning can move beyond listening and memorising. This assumption is attractive. It is also incomplete.

Class-size research has long shown that reducing class size alone does not guarantee strong learning gains. Systematic evidence suggests that small classes may support achievement, but the effects are often modest and depend on how teaching is organised [1]. A small class may create space for discussion, questioning, and feedback. It may also become a smaller lecture if the instructor continues to teach in a one-way format. For this reason, small-class teaching should not be treated as a structural cure. It is better understood as an instructional affordance. It creates opportunities, but these opportunities must be activated through teaching design.

This study begins from that problem. It asks why some small classes support deep learning while others do not. The answer proposed here is process-based. Small-class support can matter because it changes

the quality of classroom interaction. Better interaction helps students understand feedback, ask for clarification, and use feedback in later tasks. When students act on feedback, their engagement becomes more active, reflective, and self-regulated. This engagement then supports deep learning. In this logic, small-class teaching is not the direct cause of deep learning. It is the setting in which interaction, feedback uptake, and engagement can become stronger.

This argument is linked to several strands of higher education research. Active learning research has shown that students learn more when they are required to participate in tasks and discussion rather than only listen to lectures [2]. Feedback research has moved away from the idea that feedback is simply information delivered by teachers. Feedback becomes useful only when students interpret it, value it, and use it to improve their work [3]. Student engagement research treats engagement as a multidimensional process that includes behavioural, emotional, cognitive, and agentic dimensions [4], [5]. Recent evidence maps and learning analytics reviews further show that engagement is complex and should not be reduced to a single behavioural indicator [6], [7]. Deep learning research suggests that high-quality learning involves meaning-making, integration, critique, transfer, and reflection rather than surface-level reproduction [8], [9].

These bodies of work point to the same issue. The effect of small-class teaching should be examined through the learning processes that occur inside the classroom. A class becomes educationally meaningful when students speak, listen, compare ideas, revise work, and rethink what they know. The smaller size may help, but it does not replace those processes.

The present study develops and tests a sequential mediation model: small-class instructional support is expected to predict interaction quality; interaction quality is expected to predict feedback uptake; feedback uptake is expected to predict student engagement; and student engagement is expected to predict deep learning. This chain is not a statistical convenience. It follows the temporal and pedagogical order of a course. Students first experience the small-class environment. They then participate in interaction and receive feedback. They decide whether and how to use feedback. Their engagement changes through this process. Deep learning appears at the end as a learning outcome.

The study uses a three-wave mixed-methods design. Survey data were collected from 612 undergraduate students across four universities. The design also includes classroom observations, assignment-revision coding, and interviews with students and instructors. These additional data sources are important. They reduce the risk of relying only on self-reported perception. They also show whether feedback uptake and deep learning are visible in classroom events and student work.

This study makes three contributions. First, it reframes small-class teaching as a conditional affordance rather than a direct treatment. This helps explain why small classes may produce uneven effects across courses. Second, it brings feedback uptake into the analysis of small-class pedagogy. Many studies examine feedback provision, but fewer examine whether students actually use feedback. Third, it connects classroom process data with structural equation modeling. The study therefore offers both a

statistical test of the mediation model and qualitative evidence about how the mechanism works in practice.

2. Literature Review and Hypothesis Development

2.1 Small-class teaching as a conditional affordance

Small-class teaching is often linked to better access to teachers, more opportunities to speak, closer peer relations, and more flexible classroom management. These advantages are plausible. A small class allows an instructor to notice students' confusion more easily. It also makes it easier to organise discussion, give targeted feedback, and adapt tasks to student responses. Yet these are possibilities, not guaranteed outcomes.

Class-size research has warned against treating smaller classes as automatically effective. Evidence from systematic reviews suggests that class-size reduction may produce benefits, but those benefits are not uniform and may depend on teaching quality, student characteristics, and classroom organisation [1]. This makes it necessary to separate class size from small-class pedagogy. Class size is a structural feature. Small-class pedagogy is a set of enacted practices.

This study uses the concept of small-class instructional support. It refers to students' perceived access to teacher attention, participation opportunities, timely help, adaptive teaching, and individual guidance in a small-class environment. This construct is different from a simple count of students. It focuses on how students experience the instructional affordances created by the smaller setting. A class with 25 students may still have weak support if the teacher lectures throughout. A class with strong support provides opportunities for students to speak, ask questions, receive feedback, and revise their understanding.

From this view, small-class instructional support should first affect interaction quality. The smaller setting gives teachers and students more chances to build dialogue. Students can ask questions with lower social cost. Teachers can respond with more detail. Peers can hear and build on one another's ideas. These features make interaction less transactional and more dialogic.

H1. Small-class instructional support is positively associated with interaction quality.

2.2 Interaction quality: from classroom talk to dialogic learning

Interaction quality is not the same as interaction frequency. A classroom may contain many exchanges but still remain shallow. High-quality interaction involves explanation, questioning, uptake of ideas, comparison of viewpoints, and joint construction of meaning. Dialogic accounts of learning emphasise that knowledge is developed through talk, response, and shared reasoning rather than simple transmission [10].

In undergraduate classrooms, interaction quality matters because it helps students make their thinking visible. When students explain their reasoning, instructors can identify misconceptions. When students respond to peers, they learn to compare claims and evidence. When teachers ask open questions, students are pushed to articulate assumptions rather than only recall information. These interactional moves are especially important for deep learning because they require students to connect, evaluate, and reorganise knowledge.

Small-class settings can support this kind of interaction. The teacher can manage discussion more carefully. Students can receive more direct responses. The class can move from single-turn question-answer patterns to multi-turn dialogue. Interaction quality therefore becomes the first mechanism through which small-class support can matter.

2.3 Feedback uptake as the link between interaction and learning action

Feedback is central to learning, but feedback does not work simply because it is given. Research on feedback literacy has made this point clear. Students need to understand feedback, judge its value, manage emotional responses, and use feedback to improve learning or future work [3]. A systematic review of feedback reciprocity also shows that learners' active engagement with feedback is often the missing link between feedback provision and improvement [11].

This study uses the term feedback uptake to describe students' active use of feedback. It includes reading or listening to feedback carefully, asking for clarification, revising assignments, changing learning strategies, and transferring feedback to later tasks. Feedback uptake is stronger when students can discuss feedback with teachers or peers. It is weaker when feedback is vague, late, difficult to interpret, or emotionally threatening.

Interaction quality should predict feedback uptake. In a dialogic classroom, feedback is less likely to remain a private written comment. Students can ask what the comment means. Teachers can explain standards. Peers can compare interpretations. Students can see how others revise their ideas. This helps feedback become usable. Classic feedback research shows that effective feedback clarifies goals, current performance, and next steps [12]. Recent research on dialogic and peer feedback also shows that uptake improves when feedback is embedded in a social process of explanation, response, and revision [13]-[15].

H2. Interaction quality is positively associated with feedback uptake.

2.4 Student engagement as the process that turns feedback into sustained learning

Student engagement has become a central concept in higher education research. It is usually understood as a multidimensional construct that includes behaviour, emotion, cognition, and agency [4], [5]. Behavioural engagement includes attendance, effort, participation, and task completion. Emotional

engagement includes interest, belonging, and positive attachment to learning. Cognitive engagement includes deep strategy use, self-regulation, and intellectual effort. Agentic engagement includes students' proactive contribution to the learning process, such as asking questions, seeking help, and shaping classroom activity.

Feedback uptake should increase student engagement because using feedback requires action. A student who revises work after feedback is not only receiving information. The student is comparing current work with expected standards, identifying gaps, making decisions, and investing effort in improvement. These behaviours are forms of engagement. They can also make engagement more cognitive and agentic, not only behavioural.

The ICAP framework provides a useful theoretical bridge here. It argues that learning activities differ in the level of cognitive engagement they generate, from passive reception to active manipulation, constructive generation, and interactive co-construction [9]. Feedback uptake can move students upward along this continuum. When students only read feedback, engagement may remain limited. When they use feedback to revise, justify, and reapply ideas, engagement becomes constructive and interactive.

H3. Feedback uptake is positively associated with student engagement.

2.5 Deep learning as the outcome of interaction, feedback, and engagement

Deep learning refers to meaning-oriented learning. It involves understanding underlying ideas, connecting knowledge, questioning assumptions, applying concepts to new situations, and reflecting on changes in understanding. This tradition has roots in research on students' approaches to learning [8]. More recent reviews show that deep approaches do not simply increase with time in higher education; they depend on teaching, assessment, and the learning environment [16].

Student engagement should predict deep learning because engaged students are more likely to invest effort in sense-making rather than reproduction. They participate in discussion, respond to feedback, revise work, and connect ideas. These actions are the practical conditions for deep learning. A student cannot easily develop deep understanding while remaining passive.

The model proposed in this study is sequential. Small-class instructional support first creates the possibility of stronger interaction. Interaction quality helps students interpret and use feedback. Feedback uptake increases engagement. Engagement supports deep learning. The logic is ordered but not mechanical. Each link needs pedagogical work. This is why the model can explain why some small classes work and others do not. Figure 1 summarises this proposed mechanism and visually locates the direct path within the sequential mediation chain.

H4. Student engagement is positively associated with deep learning.

H5. Interaction quality, feedback uptake, and student engagement sequentially mediate the relationship between small-class instructional support and deep learning.

H6. Small-class instructional support has a positive direct association with deep learning after the mediating variables are controlled.

3. Method

3.1 Research design

This study used a three-wave mixed-methods design. The design was chosen for two reasons. First, the proposed mechanism has a temporal order. Students first experience small-class instructional support. They then participate in classroom interaction and receive feedback. They decide whether to use feedback. Their engagement develops through the course. Deep learning is expected to appear later as a learning outcome. A single-wave survey would not fit this logic well. Second, the study examines both perception and practice. Survey data were therefore combined with classroom observations, assignment-revision coding, and interviews.

The quantitative component tested the measurement model and the sequential mediation model. The qualitative and behavioural components were used to examine whether the statistical pathway was visible in classroom events, assignment revisions, and participants' explanations. This design reduces reliance on one data source and gives a stronger account of the mechanism.

Survey data were collected in three waves. At Time 1, students reported small-class instructional support and background information. At Time 2, students reported interaction quality and feedback uptake. At Time 3, students reported student engagement and deep learning. Classroom observations were conducted during the middle part of the semester, when discussion and feedback activities were most visible. Assignment-revision data were collected from courses that required draft-feedback-revision cycles. Interviews were conducted after Time 3, when students and instructors could reflect on the whole course. Figure 2 summarises the three-wave timing and the integration of survey, observation, assignment-revision, and interview evidence.

3.2 Participants and setting

The final empirical dataset included 612 undergraduate students from four Chinese universities, 38 classes, and 27 courses. Of the students, 347 were female and 265 were male. The sample included first-year, second-year, third-year, and fourth-year students. The disciplinary distribution covered education, business and management, literature and communication, social sciences, and STEM general courses. There were 368 students in small classes with no more than 30 students and 244 students in regular classes with 40 to 80 students. The regular-class group was used for group comparison and robustness checks rather than as the main theoretical focus.

The sample was suitable for the analysis for three reasons. First, students were nested in real course sections, which allowed the study to test whether the findings held after classroom clustering was considered. Second, the sample covered both small-class and regular-class settings, which made it possible to examine whether the proposed mechanism was stronger in small classes. Third, the courses included a range of disciplines, which reduced the risk that the findings only reflected one specific course type.

The continuous control variables included prior GPA, actual class size, teacher experience, and perceived course difficulty. These variables were included because student learning can be shaped by academic preparation, class size, instructor experience, and course demands. Controlling these factors made the test of the proposed mechanism more conservative.

3.3 Measures

All student survey items used a five-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree. The measures were adapted from established research on learning engagement, feedback literacy, and deep approaches to learning, with wording adjusted to the small-class context. The questionnaire was reviewed by two higher education researchers and three instructors experienced in small-class teaching. A pilot test was conducted before the main study to check item clarity.

Small-class instructional support measured students' perceived access to individual attention, opportunities to ask questions, timely help, adaptive teaching, and classroom support. Interaction quality measured dialogic exchange, reasoning, peer response, respectful discussion, and cognitive challenge. Feedback uptake measured students' use of teacher or peer feedback in revision, clarification, learning strategy adjustment, and later tasks. Student engagement measured behavioural, emotional, cognitive, and agentic engagement. Deep learning measured meaning-making, knowledge integration, critical evaluation, transfer, reflection, and learning beyond memorisation.

The five constructs were treated as latent variables. Reliability was assessed using Cronbach's alpha and composite reliability. Convergent validity was assessed using factor loadings and average variance extracted. Discriminant validity was assessed using the HTMT criterion [17].

3.4 Classroom observation and assignment-revision coding

Classroom observation was used to examine whether the survey-based constructs corresponded to actual classroom practices. Twenty class sessions were observed, including 16 small-class sessions and four regular-class sessions. The observation rubric covered interaction quality, feedback quality, balanced student participation, peer collaboration, and visible signs of deep learning. Each dimension was scored from 1 to 5. The observers also recorded classroom event frequencies, including open-ended questions, multi-turn discussions, targeted feedback events, and student clarification behaviours.

Assignment-revision coding was used to provide behavioural evidence for feedback uptake. A total of 217 student assignments were collected from courses that required at least one draft-feedback-revision cycle. Each case included the initial draft, teacher feedback, revised version, and, where available, a short student reflection. Feedback uptake was coded from 0 to 4: 0 = no uptake, 1 = surface uptake, 2 = partial uptake, 3 = deep uptake, and 4 = transfer uptake. Revision depth was coded using indicators of conceptual clarification, argument restructuring, evidence integration, critical reflection, and application to new tasks.

To strengthen coding reliability, two trained coders independently coded a randomly selected subset of the classroom observation records and assignment-revision cases. The subset included at least 25% of the observation records and 25% of the assignment-revision cases. Coding disagreements were discussed after independent coding, and the codebook was refined before full coding. Inter-coder reliability was assessed using weighted agreement for feedback-uptake levels and intraclass agreement for continuous observation and revision-depth scores. The reliability outputs were retained with the supplementary coding materials.

3.5 Qualitative interviews and coding

Semi-structured interviews were conducted with 28 students and 10 instructors. Student participants were selected to represent different levels of engagement, disciplines, and class types. Instructor participants were selected from courses included in the observation and survey data. Student interviews focused on classroom participation, feedback interpretation, feedback use, assignment revision, and perceived deep learning. Instructor interviews focused on small-class design, interaction management, feedback practices, and perceived differences between small and regular classes.

The interview analysis followed a thematic approach [18]. Initial coding identified recurring statements about visibility, dialogue, feedback use, revision, participation pressure, and deep learning. Focused coding then grouped these codes into larger themes. The qualitative analysis was not used to create new statistical variables. It was used to explain how participants understood and enacted the pathway tested in the SEM.

3.6 Quantitative analysis procedure

The quantitative analysis followed six steps. First, data screening was conducted. Missing values, outliers, response patterns, and normality were examined. Students with invalid response patterns were excluded before the final analysis. The three-wave data were matched using anonymised identifiers. The final sample included students with valid data on the main study variables. Figure 4 provides an overview of the reproducibility workflow linking data preparation, model testing, and triangulation.

Second, descriptive statistics and correlations were calculated. Third, confirmatory factor analysis was used to test the five-factor measurement model. Model fit was evaluated using chi-square divided by degrees of freedom, CFI, TLI, RMSEA, and SRMR. The fit criteria followed common SEM practice, while avoiding overreliance on any single cutoff [19], [20]. Fourth, structural equation modeling was used to

test the hypothesised paths. Fifth, bootstrap mediation analysis was conducted with 5,000 resamples because indirect effects often have non-normal sampling distributions [21]. Sixth, robustness checks were conducted. These included multilevel SEM, small-class versus regular-class comparison, alternative outcome variables, exclusion of extreme class sizes, discipline subgroup analysis, exclusion of extreme GPA groups, and comparison with PLS-SEM results. Multilevel SEM was used because students were nested within classes, and ignoring such clustering can bias standard errors and indirect-effect estimates [22], [23].

3.7 Model specification

$$IQ_i = \alpha_1 + \beta_1 SCS_i + \gamma_1 C_i + \varepsilon_{1i}$$

$$FU_i = \alpha_2 + \beta_2 IQ_i + \beta_3 SCS_i + \gamma_2 C_i + \varepsilon_{2i}$$

$$SE_i = \alpha_3 + \beta_4 FU_i + \beta_5 IQ_i + \beta_6 SCS_i + \gamma_3 C_i + \varepsilon_{3i}$$

$$DL_i = \alpha_4 + \beta_7 SE_i + \beta_8 FU_i + \beta_9 IQ_i + \beta_{10} SCS_i + \gamma_4 C_i + \varepsilon_{4i}$$

$$IE = \beta_1 \times \beta_2 \times \beta_4 \times \beta_7$$

$$TE = DE + \Sigma IE_m$$

The model does not claim experimental causality. It tests a theoretically ordered longitudinal association. The three-wave design strengthens temporal ordering, but the interpretation remains mechanism-based rather than causal in the strict experimental sense.

4. Results

4.1 Sample profile

The final empirical dataset included 612 undergraduate students from four universities, 38 classes, and 27 courses. The sample contained 368 students from small classes and 244 students from regular classes. Table 1 reports the demographic and course-level characteristics, and Table 2 reports the continuous control variables used in the models.

Table 1
Sample characteristics

Variable	Category	n	%
Gender	Female	347	56.7
Gender	Male	265	43.3
Year level	First year	168	27.45
Year level	Second year	192	31.37
Year level	Third year	176	28.76
Year level	Fourth year	76	12.42
Discipline	Education	153	25
Discipline	Business & Management	122	19.93
Discipline	Literature & Communication	126	20.59
Discipline	Social Science	118	19.28
Discipline	STEM General Courses	93	15.2
Class type	Small class (< = 30)	368	60.13
Class type	Regular class (40–80)	244	39.87
Course type	Compulsory	401	65.52
Course type	Elective	211	34.48

Table 2
Continuous control variables

Variable	M	SD	Minimum	Maximum
prior_gpa	3.16	0.57	1.82	4.89
actual_class_size	28.74	11.32	16	76
teacher_experience	8.41	5.26	1	26
course_difficulty	3.23	0.78	1	5

4.2 Measurement model

The five-factor CFA showed strong reliability and validity. All constructs exceeded the conventional reliability threshold of .70. AVE values were above .50, and factor loadings were in acceptable ranges. The five-factor measurement model showed excellent fit: $\chi^2/df = 1.723$, RMSEA = .035, CFI = .968, TLI = .962, and SRMR = .031. Table 3 presents the reliability, CR, AVE, and factor-loading results for the five constructs.

These results indicate that the five constructs were empirically distinguishable. The HTMT values were below the recommended threshold, suggesting adequate discriminant validity. The common-method assessment also suggested that same-source bias was not severe enough to account for the observed relationships [24]. This point is important because the study combines survey data with observation, assignment-revision coding, and interview evidence rather than relying on a single self-report source.

Table 3
Reliability and convergent validity

Construct	Items	Cronbach α	CR	AVE	Factor loading range
Small-class instructional support (SCS)	6	0.872	0.895	0.593	0.712–0.846
Interaction quality (IQ)	6	0.891	0.91	0.628	0.740–0.861
Feedback uptake (FU)	6	0.885	0.904	0.612	0.723–0.854
Student engagement (SE)	6	0.903	0.921	0.66	0.758–0.879
Deep learning (DL)	6	0.897	0.916	0.645	0.746–0.872

4.3 Structural equation model

The structural equation model tested the theoretically ordered pathway from small-class instructional support to deep learning through interaction quality, feedback uptake, and student engagement. The model included the control variables reported in the Method section. These controls were retained to reduce the possibility that the observed associations were driven by students prior academic preparation, course features, class size, or instructor experience.

The structural model showed good fit: $\chi^2/df = 1.816$, RMSEA = 0.037, CFI = 0.961, TLI = 0.955, and SRMR = 0.034. The hypothesised adjacent paths were all significant. Small-class instructional support was positively associated with interaction quality, $\beta = 0.527$, SE = 0.041, $p < 0.001$, 95% CI [0.446, 0.608]. Interaction quality was positively associated with feedback uptake, $\beta = 0.483$, SE = 0.044, $p < 0.001$, 95% CI [0.397, 0.569]. Feedback uptake was positively associated with student engagement, $\beta = 0.541$, SE = 0.039, $p < 0.001$, 95% CI [0.465, 0.617]. Student engagement was positively associated with deep learning, $\beta = 0.605$, SE = 0.036, $p < 0.001$, 95% CI [0.534, 0.676]. The direct path from small-class instructional support to deep learning was positive but modest, $\beta = 0.116$, SE = 0.047, $p = 0.013$, 95% CI [0.024, 0.208]. Table 4 reports the corresponding standardized structural path estimates.

The small direct association is important for interpretation. It suggests that small-class support was not mainly linked to deep learning as a simple structural condition. Its association with deep learning was largely transmitted through classroom process variables. This pattern is consistent with the theoretical claim that small classes work when they are used to build high-quality interaction, make feedback usable, and sustain student engagement.

Table 4
Standardized structural path estimates

Path	Std. β	SE	p	95% CI	Hypothesis	Conclusion
SCS -> IQ	0.527	0.041	< 0.001	[0.446, 0.608]	H1	Supported
IQ -> FU	0.483	0.044	< 0.001	[0.397, 0.569]	H2	Supported
FU -> SE	0.541	0.039	< 0.001	[0.465, 0.617]	H3	Supported
SE -> DL	0.605	0.036	< 0.001	[0.534, 0.676]	H4	Supported
SCS -> DL	0.116	0.047	0.013	[0.024, 0.208]	H6	Supported

4.4 Bootstrap mediation analysis

Bootstrap mediation analysis was conducted with 5,000 resamples. To keep the mediation test consistent with the hypothesised structural model, the analysis focused on the theoretically specified sequential pathway: small-class instructional support -> interaction quality -> feedback uptake -> student engagement -> deep learning. Non-hypothesised shortcut paths were not interpreted as independent mediation paths because they were not part of the adjacent-process model.

The sequential indirect effect was significant, $\beta = 0.083$, Boot SE = 0.021, 95% CI [0.042, 0.124]. This result supports H5. The finding means that small-class instructional support was associated with deep learning through a time-ordered process: students first experienced stronger interaction quality, then reported higher feedback uptake, then showed stronger engagement, and finally reported deeper learning. Table 5 reports the bootstrap estimate for the hypothesised sequential indirect effect, and Fig. 3 visualizes the indirect and direct effects.

The indirect effect was not interpreted as proof of experimental causality. It was interpreted as evidence for a theoretically ordered mechanism supported by three-wave data and supplementary classroom-process evidence. This interpretation is more appropriate because students were nested in real courses and were not randomly assigned to all instructional conditions.

Table 5
Bootstrap test of the hypothesised sequential indirect effect

Indirect pathway	Standardized effect	Boot SE	95% CI	Conclusion
SCS -> IQ -> FU -> SE -> DL	0.083	0.021	[0.042, 0.124]	Significant
Direct effect: SCS -> DL	0.116	0.047	[0.024, 0.208]	Significant but modest

4.5 Multilevel SEM and group comparison

Because students were nested within classes, class-level clustering was examined. The ICC values were above .12 for all major constructs: SCS = .214, IQ = .187, FU = .163, SE = .201, and DL = .175. These values justified a multilevel robustness check. In the multilevel model, the sequential indirect effect remained significant, $\beta = .079$, 95% CI [.038, .120], $p < .001$. This result suggests that the proposed mechanism was not merely an artefact of ignoring the nested structure of the data.

The multilevel result should be interpreted as a robustness test rather than as a separate causal model. Its purpose was to examine whether the interaction-feedback-engagement pathway remained stable after shared class-level variance was considered. The consistency between the single-level SEM and multilevel SEM strengthens confidence in the proposed mechanism. Table 6 shows the latent mean comparison between small and regular classes.

Table 6
Latent mean comparison between small and regular classes

Construct	Small class (N = 368)	Regular class (N = 244)	Group difference p
SCS	3.87	2.61	< 0.001
IQ	3.72	2.48	< 0.001
FU	3.64	2.53	< 0.001
SE	3.79	2.7	< 0.001
DL	3.75	2.66	< 0.001

In the small-class group, the sequential mediation effect was .112 and significant. In the regular-class group, the same chain was .027 and the confidence interval included zero. The finding suggests that the full interaction-feedback-engagement-learning pathway was more strongly activated in the small-class setting.

4.6 Classroom observation evidence

Classroom observation was used as supplementary process evidence. Because the number of observed regular-class sessions was limited, the observation results were not treated as an independent causal test. They were used to examine whether the mechanism suggested by the survey model could also be seen in classroom practices.

Across the observed sessions, small classes showed stronger process indicators than regular classes. The mean score for interaction quality was 4.12 in small classes and 2.35 in regular classes. Feedback quality was also higher in small classes, with mean scores of 3.96 and 2.18 respectively. Balanced participation, peer collaboration, and visible signs of deep learning showed the same pattern. The descriptive effect sizes were large, but they should be read cautiously because the regular-class observation sample was small.

Event-frequency records supported the same interpretation. Small classes had more open-ended questions, more multi-turn discussions, more targeted feedback events, and more student clarification behaviours. On average, small classes included 12.7 open-ended questions per session, compared with 3.2 in regular classes. Targeted feedback events occurred 9.6 times per small-class session and 2.5 times per regular-class session. Student clarification behaviours were also more frequent in small classes.

These observation findings are consistent with the quantitative model. They suggest that small classes created more visible opportunities for dialogue and feedback use. The observation evidence does not prove the mechanism by itself. It shows that the statistical pathway was grounded in observable classroom processes. Table 7 summarises the observation ratings, and Table 8 reports the corresponding event-frequency records.

Table 7
Classroom observation results

Observation dimension	Small classes M $\pm$ SD	Regular classes M $\pm$ SD	Descriptive interpretation
Interaction quality	4.12 $\pm$ 0.47	2.35 $\pm$ 0.52	Higher in small classes
Feedback quality	3.96 $\pm$ 0.51	2.18 $\pm$ 0.44	Higher in small classes
Balanced student participation	3.84 $\pm$ 0.58	1.76 $\pm$ 0.63	Higher in small classes
Peer collaboration	3.71 $\pm$ 0.62	1.92 $\pm$ 0.57	Higher in small classes
Visible signs of deep learning	3.88 $\pm$ 0.54	2.04 $\pm$ 0.49	Higher in small classes

Table 8
Classroom event-frequency records

Classroom event	Small classes: mean per session	Regular classes: mean per session
Open-ended teacher questions	12.7	3.2
Multi-turn dialogic exchanges	8.4	1.1
Targeted feedback events	9.6	2.5
Student clarification behaviours	7.2	0.8

4.7 Assignment-revision and qualitative evidence

Assignment-revision coding provided behavioural evidence for feedback uptake. Among the 217 coded assignments, 29 cases showed no uptake of feedback, accounting for 13.36% of the sample. Surface

uptake appeared in 46 cases, accounting for 21.20%. Partial uptake was the most common category, with 74 cases, or 34.10%. Deep uptake appeared in 51 cases, accounting for 23.50%. Transfer uptake appeared in 17 cases, accounting for 7.84%. Table 9 reports the full distribution of feedback-uptake coding levels.

This distribution is useful because it shows that feedback uptake was not uniformly high. Many students used feedback only partially or at a surface level. This makes the evidence more credible than a uniformly positive pattern. It also fits the theoretical claim that feedback uptake is a learning process that varies across students.

The coded feedback-uptake score was positively correlated with the survey-based feedback-uptake scale, $r = .642$, $p < .001$. Revision-depth coding was positively correlated with the deep-learning scale, $r = .587$, $p < .001$. These correlations suggest that students self-reported feedback uptake and deep learning were reflected in their written revisions. Table 10 reports the behavioural validation correlations. The assignment evidence therefore strengthens the interpretation of feedback uptake as a behavioural process rather than only a favourable perception.

Table 9
Feedback-uptake coding results

Coding level	Meaning	n	%
0	No uptake	29	13.36
1	Surface uptake	46	21.2
2	Partial uptake	74	34.1
3	Deep uptake	51	23.5
4	Transfer uptake	17	7.84

Table 10
Behavioural validation of feedback uptake

Relationship	r	p	Interpretation
Coded feedback uptake ↔ survey feedback uptake	0.642	< 0.001	Strong positive correspondence
Revision depth ↔ survey deep learning	0.587	< 0.001	Moderate-to-strong positive correspondence

The interview evidence explained the process behind these results. Students often described small classes as spaces where they felt more visible and more able to ask follow-up questions. This visibility did not automatically lead to participation. Some students felt more pressure in small classes. Yet when instructors used structured questioning and feedback dialogue, students were more likely to clarify comments and revise their work.

Three themes were identified. First, small classes increased students visibility and made dialogic exchange more likely. Second, feedback became useful when students could ask what it meant and how it should be acted on. Third, sustained engagement developed through repeated cycles of discussion, feedback, revision, and reflection. Instructors gave a similar account. They emphasised that small classes were valuable only when the teacher designed tasks, managed participation, and gave feedback that students could actually use.

The qualitative findings therefore support the proposed mechanism without simply repeating the survey results. They show how interaction quality helped students make sense of feedback, how feedback uptake became visible in assignment revision, and how repeated revision supported deeper learning.

4.8 Robustness checks

Several robustness checks were conducted. First, course gain and learning satisfaction were used as alternative outcome variables. The sequential pathway remained positive and significant, although the effect size was smaller than that for deep learning. Second, students in extreme class-size groups were excluded. The main path coefficients changed by less than 0.04, suggesting that the findings were not driven by unusually small or unusually large classes. Third, discipline subgroup analyses were conducted. The pathway remained positive across the five disciplinary groups, although the size of the indirect effect varied.

Fourth, the model was re-estimated after excluding students in the top and bottom 10% of prior GPA. The direction and significance of the main pathway remained stable. Fifth, a PLS-SEM specification was estimated as an additional robustness check. The main pattern was consistent with the covariance-based SEM results. Sixth, the multilevel SEM addressed the nested structure of students within classes.

These checks were not used to claim that the model is the only possible explanation. They were used to examine whether the proposed mechanism remained stable under alternative specifications. Overall, the results support the claim that the relationship between small-class support and deep learning is mainly carried by interaction quality, feedback uptake, and student engagement.

5. Discussion

5.1 Overview of the findings

This study examined how small-class instructional support is associated with deep learning in undergraduate classrooms. The results support the proposed mechanism. Small-class support was linked to stronger interaction quality. Better interaction predicted higher feedback uptake. Feedback uptake predicted stronger student engagement. Engagement then predicted deep learning. The direct path from small-class support to deep learning was positive but modest.

This pattern is important for the central argument of the article. It shows that small-class teaching should not be treated as a simple class-size effect. Students did not benefit from the smaller class

simply because fewer students were present. They benefited when the smaller setting supported dialogue, feedback use, and sustained engagement. The observation data and assignment-revision evidence were consistent with this interpretation. Small classes included more open-ended questions, more targeted feedback events, and more student clarification behaviours. Assignment coding also showed that students who reported stronger feedback uptake tended to make deeper revisions.

The findings therefore shift the focus from class size to classroom process. A small class is not automatically a deep-learning environment. It becomes one when the instructor uses the smaller setting to organise meaningful interaction, make feedback actionable, and involve students in revision and reflection.

5.2 Theoretical contribution: from class size to instructional affordance

The first theoretical contribution is the reframing of small-class teaching as an instructional affordance. Much class-size research asks whether smaller classes improve outcomes. That question is useful, but it is incomplete. A class can be small and still remain teacher-centred. A small class can also produce silence, pressure, or passive compliance. The educational value of a small class depends on how its affordances are used.

This study shows that small-class support works through process variables. Interaction quality is the first link. The smaller setting gives students and teachers more chances to ask, respond, explain, and revise ideas. Yet this opportunity has to be designed. Without open questions, discussion tasks, and teacher responsiveness, the class size itself does little.

This contribution matters for Q1-level higher education research because it avoids a simple policy message. The study does not say that universities should only reduce class size. It says that class-size reform needs a pedagogical mechanism. Smaller classes should be evaluated by what they make possible: dialogue, feedback uptake, engagement, and deep learning.

5.3 Feedback uptake as the central mechanism

The second contribution is the use of feedback uptake as a central mechanism. Feedback is often treated as something teachers provide. This study treats feedback as something students do. Feedback becomes educationally useful when students understand it, ask about it, revise with it, and transfer it to later tasks.

The assignment-revision data are important here. They show that feedback uptake was visible in student work. Not all students used feedback deeply. Some ignored it. Some made surface-level changes. Others restructured arguments or transferred feedback to new tasks. This variation gives the mechanism more credibility. It shows that feedback uptake is not just a positive survey response. It is a behavioural process with different levels of depth.

This point also explains why interaction quality matters. Students are more likely to use feedback when they can discuss it. In a dialogic classroom, feedback is not a final judgement. It becomes part of an ongoing exchange. Students can ask what the feedback means, compare it with criteria, and test possible revisions. Small classes can support this process because teachers can follow students thinking more closely.

5.4 Engagement as the bridge to deep learning

The third contribution concerns student engagement. Engagement is sometimes used as a broad indicator of student involvement. In this study, engagement has a more precise role. It is the bridge between feedback uptake and deep learning.

Students who take up feedback do more than correct errors. They monitor their work, compare standards, rethink ideas, and invest effort in revision. These actions are forms of behavioural, cognitive, emotional, and agentic engagement. They require students to become active participants in learning rather than passive recipients of teaching.

This helps explain why engagement predicts deep learning. Deep learning requires meaning-making, integration, critique, and transfer. These processes rarely appear without sustained engagement. The results support this logic. Feedback uptake predicted engagement, and engagement predicted deep learning. The finding suggests that deep learning grows out of repeated learning actions, not from class size alone.

5.5 Methodological contribution

The study also makes a methodological contribution. Many classroom studies rely only on a one-time questionnaire. This article used three-wave survey data, classroom observation, assignment-revision coding, interviews, and multilevel SEM. The design is not experimental, but it is stronger than a single cross-sectional model.

The three-wave design fits the theoretical order of the mechanism. Small-class support was measured earlier. Interaction and feedback were measured during the course. Engagement and deep learning were measured later. This temporal structure does not prove causality, but it reduces one common weakness in mediation research: measuring all variables at the same time and then imposing a causal order.

The observation and assignment evidence also improve credibility. They show whether the mechanism appears outside self-reported perception. The observation data show process differences in classroom practice. The assignment data show feedback uptake in written revision. The interviews explain why some students used feedback deeply while others remained at the surface level.

5.6 Causal caution and interpretation

The findings should be interpreted carefully. The study supports a theoretically ordered mechanism, not a strict experimental causal effect. Students were not randomly assigned to all course formats. Teachers

and courses differed in design, discipline, assessment, and feedback routines. These differences may affect the results.

The article therefore avoids strong causal claims. It does not argue that small-class teaching automatically causes deep learning. It argues that small-class instructional support is associated with deep learning mainly through interaction quality, feedback uptake, and student engagement. The three-wave design, multilevel SEM, classroom observation, and assignment-revision evidence make this interpretation stronger, but they do not remove all selection concerns.

This cautious interpretation is useful for publication. It protects the article from overclaiming. It also matches the theoretical argument. Small-class teaching is not a fixed treatment. It is an opportunity structure. Its effect depends on what teachers and students do with it.

5.7 Practical implications

The findings have several practical implications. First, universities should not evaluate small-class reform only by counting students. A class with fewer than 30 students may still fail if it is taught as a smaller lecture. Evaluation should include interaction quality, feedback usability, participation distribution, and evidence of student revision.

Second, instructors need support in small-class pedagogy. Teaching fewer students does not automatically produce better teaching. Teachers need to design open questions, structured discussion, peer response, feedback dialogue, and revision tasks. These practices turn the smaller setting into a learning environment.

Third, feedback should be designed as a process. Written comments are not enough. Students need time to interpret feedback, ask questions, revise, and apply suggestions to later work. Small classes are well suited to this process because teachers can follow students revisions more closely.

Fourth, participation should be structured. Some students benefit from small classes because they feel visible. Others may feel exposed. Teachers should not rely only on confident students who volunteer to speak. Short written reflections, pair discussion, rotating roles, peer explanation, and revision logs can make participation more balanced.

Fifth, assignment revision should be used as learning evidence. The study shows that revision depth is closely linked to feedback uptake and deep learning. Courses that include draft-feedback-revision cycles may be especially suitable for small-class teaching.

5.8 Limitations and future research

Several limitations should be acknowledged. First, the sample came from four Chinese universities. The courses covered different disciplines, but the findings may still reflect the institutional and cultural features of Chinese undergraduate education. Future studies should test the model in different countries and institutional types.

Second, the study used a longitudinal design but not a randomised experiment. The three-wave structure supports temporal ordering, but it cannot remove all selection effects. Students in small classes may differ from those in regular classes in motivation, academic preparation, or expectations. The study controlled for prior GPA, course type, class size, teacher experience, and course difficulty. It also used multilevel SEM. Still, future research can use matched comparison groups, natural experiments, or course-level reform rollouts.

Third, some constructs were measured by student self-report. The study reduced this concern through classroom observation, assignment-revision coding, and interviews. Future research can go further by adding video-coded classroom talk, feedback-text analysis, learning analytics, portfolio assessment, and objective performance tasks.

Fourth, the classroom observation comparison included fewer regular-class sessions than small-class sessions. The observation evidence was therefore used as supplementary process evidence, not as an independent causal test. Future studies should include a larger and more balanced observation sample.

Fifth, the assignment-revision data came from courses that used draft-feedback-revision cycles. These courses may already have stronger feedback cultures. Future studies can compare courses with and without structured revision requirements to examine whether feedback uptake depends on assessment design.

Sixth, deep learning was measured through survey responses and revision evidence. Deep learning is complex. Future work can use concept maps, transfer tasks, oral defence, portfolio assessment, or discipline-specific problem-solving tasks to capture deeper forms of understanding.

Future research can also examine boundary conditions. Psychological safety, teacher immediacy, peer belonging, feedback literacy, epistemic curiosity, and self-efficacy may shape whether students participate and use feedback. These factors may be especially important for low-achieving students, first-year students, and students who are less confident in discussion-based learning.

6. Conclusion

This study examined how small-class instructional support is linked to deep learning in undergraduate classrooms. The findings show that small classes do not work simply because they are small. Their value lies in the learning processes they make possible.

The results support a sequential mechanism. Small-class support was associated with stronger interaction quality. Interaction quality was associated with higher feedback uptake. Feedback uptake was associated with stronger student engagement. Engagement was associated with deeper learning. The direct path from small-class support to deep learning was positive but modest. The main association operated through classroom process.

The study contributes to higher education research by moving the debate from class size to pedagogical mechanism. A small class should be understood as an opportunity structure. It creates better conditions for dialogue, feedback use, and engagement. It does not guarantee them. Teachers still need to design tasks, guide discussion, support feedback uptake, and help students revise their work.

For practice, the message is clear. A small class should not be a smaller lecture. It should be a feedback-rich and interaction-rich learning environment. Students need chances to speak, question, receive feedback, use feedback, revise work, and reflect on learning. When these conditions are present, small-class teaching can support deep learning. When they are absent, the advantage of small classes is likely to remain unrealised.

Declarations

Declaration of Competing Interests:

The author declares that there are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed Consent Statement:

Informed consent to participate was obtained from all individual participants included in the study.

Acknowledgements:

The author has no acknowledgements to declare.

Ethics approval and human participants

This study was approved by the Institutional Review Board of Stanford University (IRB Protocol #87656; approval date: June 26, 2026). The research was conducted in accordance with the Declaration of Helsinki and applicable institutional ethical standards.

Consent to participate

Informed consent to participate was obtained from all individual participants included in the study.

Consent to publish

Not applicable. The manuscript contains no identifiable personal data.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Funding

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

Author Contribution

Y.Z. conceived and designed the study, developed the theoretical framework, collected and analyzed the data, interpreted the findings, prepared the figures and tables, and wrote the main manuscript text. Y.Z. reviewed and approved the final manuscript and agrees to be accountable for all aspects of the work.

References

1. Filges T, Sonne-Schmidt C, Nielsen BCV. Small class sizes for improving student achievement in primary and secondary schools. *Campbell Syst Reviews*. 2018;14(1):1–107.
2. Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP. Active learning increases student performance in science, engineering, and mathematics, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 23, pp. 8410–8415, 2014.
3. Carless D, Boud D. The development of student feedback literacy: Enabling uptake of feedback, *Assessment & Evaluation in Higher Education*, vol. 43, no. 8, pp. 1315–1325, 2018.
4. Fredricks JA, Blumenfeld PC, Paris AH. School engagement: Potential of the concept, state of the evidence. *Rev Educ Res*. 2004;74(1):59–109.
5. Kahu ER, Nelson K. Student engagement in the educational interface: Understanding the mechanisms of student success, *Higher Education Research & Development*, vol. 37, no. 1, pp. 58–71, 2018.
6. Bond M, Buntins K, Bedenlier S, Zawacki-Richter O, Kerres M. Mapping research in student engagement and educational technology in higher education: A systematic evidence map, *International Journal of Educational Technology in Higher Education*, vol. 17, art. no. 2, 2020.
7. Bergdahl N, Bond M, Sjöberg J, Dougherty M, Oxley E. Unpacking student engagement in higher education learning analytics: A systematic review, *International Journal of Educational Technology in Higher Education*, vol. 21, art. no. 63, 2024.
8. Biggs JB. *Student Approaches to Learning and Studying*. Hawthorn, VIC, Australia: Australian Council for Educational Research; 1987.
9. Chi MTH, Wylie R. The ICAP framework: Linking cognitive engagement to active learning outcomes, *Educational Psychologist*, vol. 49, no. 4, pp. 219–243, 2014.
10. Mercer N, Howe C. Explaining the dialogic processes of teaching and learning: The value and potential of sociocultural theory, *Learning, Culture and Social Interaction*, 1, 1, pp. 12–21, 2012.

11. Winstone NE, Nash RA, Parker M, Rowntree J. Supporting learners agentic engagement with feedback: A systematic review and a taxonomy of recipience processes, *Educational Psychologist*, vol. 52, no. 1, pp. 17–37, 2017.
12. Hattie J, Timperley H. The power of feedback. *Rev Educ Res*. 2007;77(1):81–112.
13. Wood J. Making peer feedback work: The contribution of technology-mediated dialogic peer feedback to feedback uptake and literacy, *Assessment & Evaluation in Higher Education*, vol. 47, no. 3, pp. 327–346, 2022.
14. Little T, Dawson P, Boud D, Tai J. Can students feedback literacy be improved? A scoping review of interventions, *Assessment & Evaluation in Higher Education*, vol. 49, no. 1, pp. 39–52, 2024.
15. To J. Feedback partnerships: Strengthening students proactive recipience through co-creating dialogic feedback, *Higher Education Research & Development*, vol. 44, no. 4, 2025.
16. Asikainen H, Gijbels D. Do students develop towards more deep approaches to learning during studies? A systematic review on the development of students deep and surface approaches to learning in higher education. *Educational Psychol Rev*. 2017;29(2):205–34.
17. Henseler J, Ringle CM, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J Acad Mark Sci*. 2015;43(1):115–35.
18. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Res Psychol*. 2006;3(2):77–101.
19. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives, *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1, pp. 1–55, 1999.
20. Kline RB. *Principles and Practice of Structural Equation Modeling*. 4th ed. New York, NY, USA: Guilford Press; 2016.
21. MacKinnon DP, Lockwood CM, Williams J. Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivar Behav Res*. 2004;39(1):99–128.
22. Preacher KJ, Zyphur MJ, Zhang Z. A general multilevel SEM framework for assessing multilevel mediation. *Psychol Methods*. 2010;15(3):209–33.
23. Muthén BO. Multilevel covariance structure analysis, *Sociological Methods & Research*, vol. 22, no. 3, pp. 376–398, 1994.
24. Podsakoff PM, MacKenzie SB, Lee J-Y, Podsakoff NP. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *J Appl Psychol*. 2003;88(5):879–903.

Figures

Sequential mediation model



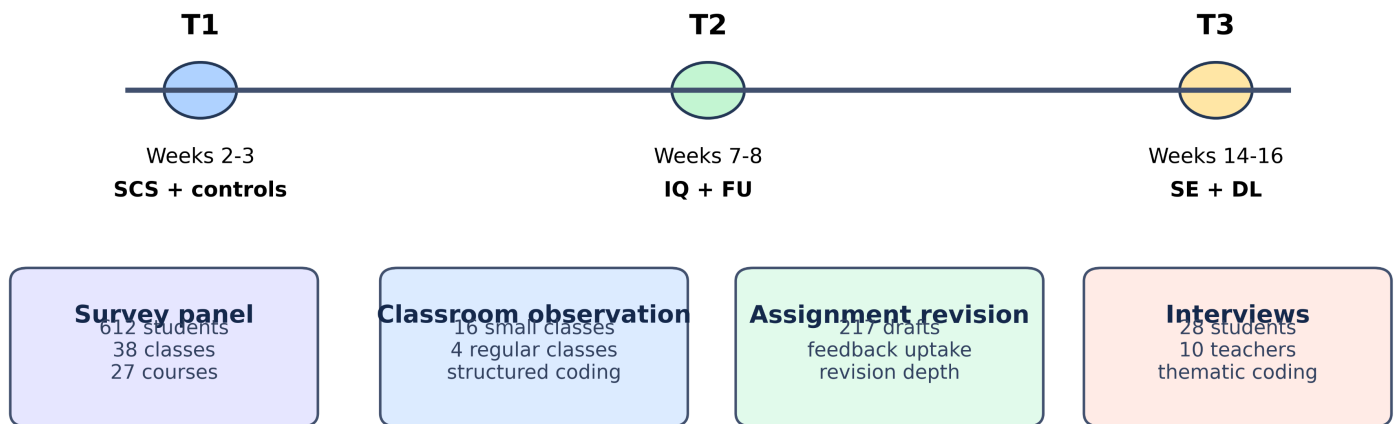
direct path retained but small: $\beta=.116$

Controls: gender, year level, discipline, prior GPA, course type, actual class size, teacher experience, course difficulty

Figure 1

Conceptual mechanism linking small-class instructional support with deep learning.

Three-wave mixed-methods data structure



Triangulation links perceived classroom processes with observed interaction, feedback use, and qualitative accounts.

Figure 2

Three-wave mixed-methods design and data integration process.

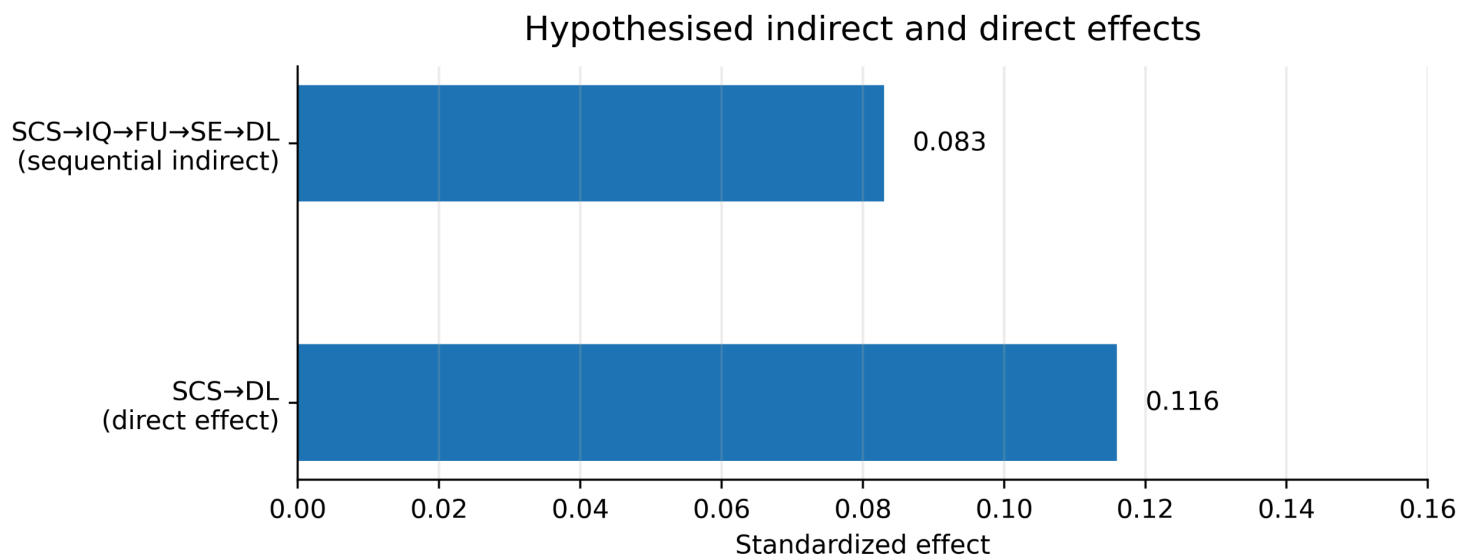


Figure 3

Hypothesised sequential indirect-effect pathway.

Analysis and reproducibility workflow

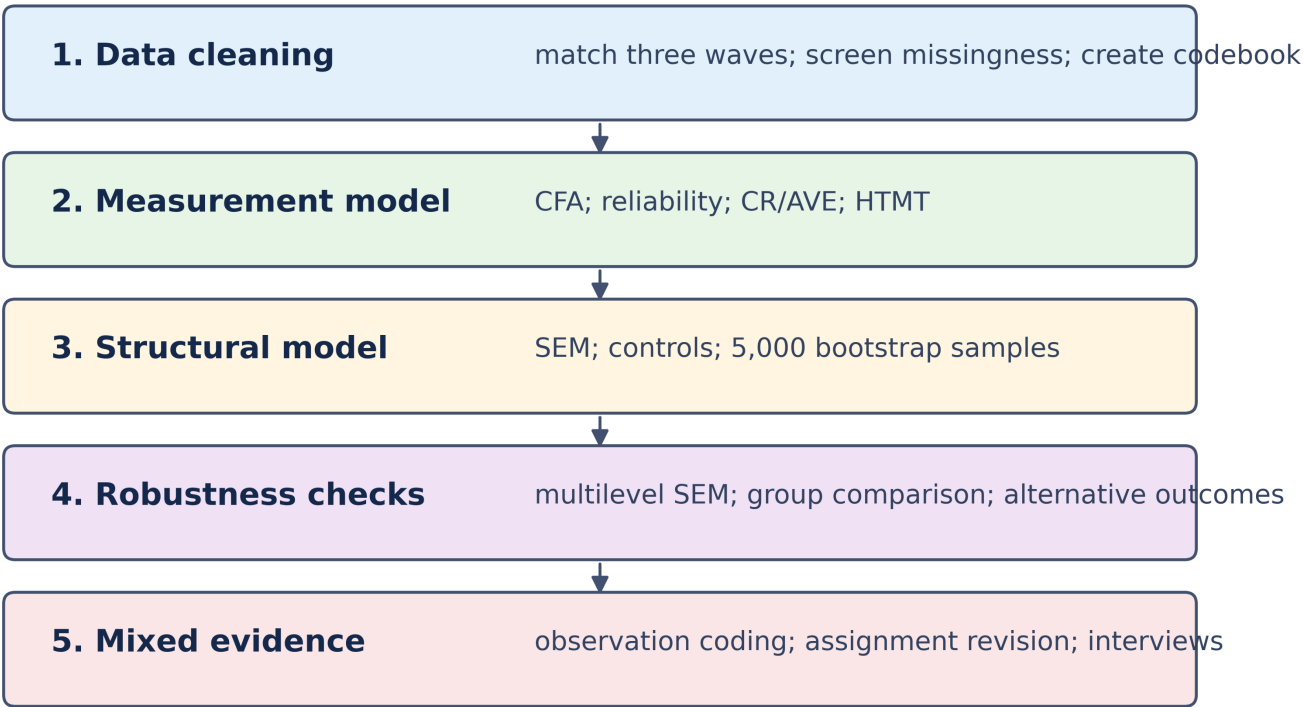


Figure 4

Reproducibility workflow for data preparation, model testing, and triangulation.