IADIS International Journal on WWW/Internet
Vol. 24, No. 2, pp. 49-63
ISSN: 1645-7641

DESIGNING FOR DIGITAL ATTENTION

Bertil P. Marques[1] and Marilío Cardoso[2]
[1]*GILT/School of Engineering of the Technical University of Porto, Porto, Portugal*
[2]*SIIS/School of Engineering of the Technical University of Porto, Porto, Portugal*

ABSTRACT

Personal devices can support learning while inviting unrelated activity. We analysed item-level records for 303 engineering students in six classes assigned to three cumulative digital lecture configurations. Reported off-task categories were lower with quizzes, a message wall and synchronised slides than with quizzes alone (3.50 versus 4.71); post-intervention engagement also differed across configurations. Yet only two classes received each configuration, assignment records were unavailable, and the data contain no baseline or repeated student-level outcomes. A class-level randomisation sensitivity analysis therefore qualifies student-level tests. The seating interaction was inconclusive. Collective feedback coincided with greater interest and comparison concern but lower participation and assessment; feedback was perfectly aligned with class, and the four-item threat scale had low internal consistency ($\alpha = .30$). The findings identify promising design contrasts while requiring documented allocation, instrument validation and independent confirmation of ethics approval before causal interpretation.

KEYWORDS

ICAP, Multitasking, Engagement, Synchronization, Feedback, Education

1. INTRODUCTION

A student's personal device can function as an immediate channel to a lecture, a private notebook, a response system and an unrelated stream of messages within the same hour. This dual role complicates the familiar question of whether devices should be allowed in the classroom. Restricting access may reduce some distractions, but it may also remove useful routes to enlarged diagrams, written questions and formative feedback. The relevant research problem concerns the design of the activity in which the device participates. What information appears on the screen, what must the student do with it, and how are their responses made visible to others? These choices affect attention, cognitive work and the social conditions for participation in different ways (Sana et al., 2013; Jamet et al., 2020).

Student engagement is commonly analysed through cognitive, affective and behavioural dimensions (Fredricks et al., 2004). The ICAP framework instead distinguishes modes of cognitive activity—passive, active, constructive and interactive—by the observable work learners undertake with information (Chi and Wylie, 2014). These approaches are complementary but cannot be substituted for one another. An enthusiastic response to a public quiz may index affect without demonstrating explanation or dialogue. Synchronising the instructor's slides to a laptop may improve access to content without causing the learner to co-construct knowledge. A theoretically informed evaluation needs to identify the task associated with each feature and measure its intended outcome separately.

Previous research warns about novelty claims. Hutain and Michinov (2022) studied three cumulative functionalities of a digital environment in a sample of 303 psychology students. In a separate study, Michinov and Hutain (2024) compared slideshow display on personal devices with a projected display and examined multitasking and affective engagement. Their designs and results cannot be treated as observations from the present engineering cohort. The current paper therefore describes an independent analysis of the workbook supplied for this project, identifies where its records support the manuscript claims, and distinguishes these records from the cited studies. A newly supplied item-level workbook permits analysis of the adapted questionnaires, while discrepancies from an earlier extract require the provenance of the revised records to be established before submission.

Three questions guide the analysis. First, are the cumulative classroom configurations associated with different counts of reported off-task activities and different engagement scores? Second, does seating position modify the multitasking association in a manner consistent with the proposed "spatial equalisation" account? Third, how do private and collective feedback correspond to interest, comparison concerns, participation and assessment? The three questions are deliberately framed as associations. The spreadsheet provides a single post-intervention observation per student; it does not contain the repeated student-level measures needed to estimate change across four sessions or to infer a longitudinal mechanism.

This distinction matters for contribution and interpretation. The work can test the internal coherence of a recorded dataset, report reproducible contrasts and generate a more precise design model. It cannot establish that occupying a screen raises the cognitive cost of task switching, or that a particular ICAP mode was enacted, without direct measures of those processes. The study is therefore positioned as a transparent analysis of classroom configurations and outcomes, coupled with a research agenda that could discriminate between competing explanations. It addresses both the technical decisions of instructors and the social consequences of showing learners how others responded.

2. THEORETICAL FRAMEWORK

2.1 Engagement and ICAP as Distinct Analytical Lenses

The construct of engagement encompasses more than attention or activity. Cognitive engagement concerns investment in understanding and regulation of learning; affective engagement concerns interest, belonging and response to the experience; behavioural engagement concerns participation and persistence (Fredricks et al., 2004). Kahu and Nelson (2018) place engagement at an educational interface shaped by students and institutional

conditions. In a digital lecture, a student can report high interest yet avoid a public response, or participate often while processing material only superficially. The three dimensions should therefore be reported individually rather than collapsed into a general claim that a platform "increased engagement".

ICAP makes a different distinction. Passive reception involves exposure without overt manipulation. Active work changes or selects from supplied information, as in answering a recognition question. Constructive work generates an explanation, question or inference that goes beyond what was supplied. Interactive work requires reciprocal construction: one participant's contribution is taken up and revised by another (Chi and Wylie, 2014). A message wall affords writing, but its actual category depends on the content of submissions and the exchanges they initiate. A slide display makes course information available but does not itself move a student to interactive processing. The original manuscript's condition labels can be retained as descriptions of intended pedagogical affordances, not as measured ICAP outcomes.

This distinction changes what evidence would count as an intervention effect. If quizzes were the active component, response logs could demonstrate participation, while explanations of answers could reveal constructive reasoning. If a wall was intended to generate questions, its posts should be coded for conceptual content and subsequent instructor or peer uptake. If synchronised slides were intended to hold attention, viewing and switching behaviour would need observation or consented logs. The workbook contains none of those trace measures. It instead offers nine engagement item responses, fourteen binary off-task indicators, four comparison-threat items and a binary participation indicator. Those are relevant outcomes, but their relationship with the theoretical mechanisms remains interpretative.

The meta-analysis by Freeman et al. (2014) offers broad evidence that active learning can improve performance in science and engineering. It should not be read as a claim that every technology labelled interactive is effective. The educational consequence of a tool depends on the reasoning it asks learners to perform, the feedback they receive and the alignment between activities and assessment. In the present comparison, the high configuration inherits quizzes and a wall while adding slides; it therefore represents a bundle of affordances. An increase in engagement across conditions cannot be attributed uniquely to slide synchronisation without comparing the incremental high and moderate configurations under equivalent teaching conditions.

2.2 Multitasking, Accessibility and the Screen

Non-academic multitasking is often described as simultaneous activity, but much classroom use involves rapid alternation between tasks. Repeated switching can impose processing costs because working memory is limited and learners must re-establish the instructional goal after attending to a competing activity (Sweller, 1988). Experimental research has found adverse consequences for both laptop multitaskers and nearby peers (Sana et al., 2013); naturalistic work illustrates the complexity of measuring actual classroom behaviour (Jamet et al., 2020). Neither literature licenses the assumption that any instructional content on a device will improve learning. The design must prevent needless division of attention without adding redundant or competing displays.

"Interface occupation" is a proposed account of how synchronised slides might matter. When the lecture appears in a student's principal workspace, maintaining contact with the material becomes easier and switching to unrelated content may be less likely. An alternative

explanation is visual access: a learner far from a projection can inspect a diagram on a nearby screen. A third is instructor behaviour: providing slides may alter the pacing or the number of opportunities for response. These mechanisms predict different observable patterns. The first requires evidence about switching frequency or application use; the second predicts a condition-by-seat or condition-by-visibility interaction; the third requires a session-level account of teaching. The workbook measures none of the mechanisms directly and records seat location only once per student.

The fourteen binary activities in the spreadsheet capture whether a category occurred, not how often it occurred or how long it lasted. The total therefore denotes breadth of reported categories, not "events per session" in the ordinary sense. One glance at a notification and an extended visit to a news website each add one point. A lower mean may be educationally useful, but it cannot be converted into a percentage reduction in attention lost or a demonstrated reduction in extraneous cognitive load. The inclusion of off-topic generative AI use among the categories is contemporary, although the dataset provides no context for deciding whether a particular AI interaction was relevant to course work.

2.3 Feedback Visibility and Social Comparison

A class-wide display can turn a poll into a conversation about reasons for divergent answers. Yet the same display can make error or minority opinion feel publicly exposed. Hutain and Michinov (2024) examined this double-edged possibility in online learning; the present workbook records a separate feedback label in a face-to-face lecture dataset. The feedback conditions must therefore be analysed independently, with the source study cited as prior literature rather than blended into the engineering results. Interest and comparison concern are not opposing ends of one scale: a student may find a public display compelling and still hesitate to speak.

Anonymity, aggregation, group size and instructional framing can change the meaning of "public". A bar chart of class responses is different from a named ranking, and a carefully framed discussion of attractive incorrect options differs from speed-based competition. The workbook specifies private versus collective feedback at participant level, but it does not show the actual displayed screens or a session-by-session assignment record. The revised workbook supplies all four item responses. Their internal consistency can now be recalculated, and low reliability would constrain the interpretation of the composite even when its group means differ.

The social dimension also places limits on subgroup interpretation. Gender is recorded in three categories, including nine participants in an "other/prefer not to say" category. A comparison with four or five people per feedback condition is unstable and raises confidentiality concerns if presented with unnecessary detail. The study therefore reports the overall feedback association and treats any gender-specific pattern as exploratory. The relevant question is whether visibility alters participation differently across learners; resolving it requires a larger sample, measured baseline anxiety and an interaction test with adequate precision.

3. METHODS

3.1 Source, Design and Unit of Allocation

The revised workbook contains seven sheets: participant characteristics, an aggregate register for four lecture sessions, nine engagement-item responses, fourteen binary multitasking indicators, four comparison-threat responses with assessment and participation fields, item wording, and a protocol narrative. Each of the four student-level sheets contains 303 unique IDs, and all IDs join one-to-one. No item responses are missing in the reported fields. This analysis takes that revised workbook as its data source, rather than blending its values with the earlier six-sheet extract. The earlier extract gives different condition assignments and different values for many students; for example, the first participant's multitasking total is four in the earlier extract but two in the revised one. A signed extraction history or original system export is needed to explain the change. Until then, every estimate in this paper is conditional on the revised file.

The revised participant sheet assigns classes A and B to quizzes, C and D to quizzes plus a message wall, and E and F to those features plus synchronised slides. Class sizes are 51 and 50 in each condition, giving 101 students per condition. This agrees with the grouping described in the revised protocol, which claims cluster randomisation. It does not supply a random sequence, allocation concealment record, class timetable, eligibility list or pre-test values. Consequently, we describe the observed class allocation without treating its randomised status or baseline equivalence as verified. The condition is constant within class; a student-level test that treats all 303 responses as independent underestimates uncertainty if classmates share an instructor, time slot or learning environment (Campbell et al., 2012).

Table 1. Observed class allocation and feedback pairing

Class	n	Configuration	Feedback
A	51	Quizzes	Private
B	50	Quizzes	Collective
C	51	Quizzes and wall	Collective
D	50	Quizzes and wall	Private
E	51	Slides added	Private
F	50	Slides added	Collective

Feedback was assigned at the class level too: one class in each configuration received private feedback and the other collective feedback (Table 1). Its association with outcomes is inseparable from six specific classes without either more clusters or within-class variation. The workbook describes four 90-minute sessions in October 2025 and reports condition-level attendance for each, but lacks student-by-session attendance, individual response timing and pre-intervention scores. The analysis therefore concerns one post-intervention record per person, not repeated measures or individual change over four lectures. Six clusters are particularly few for conventional cluster-robust standard errors or subgroup claims.

3.2 Classroom Configurations and Measures

The cumulative configurations were individual quizzes; quizzes plus an anonymous message wall; and both features plus synchronised lecture slides on students' devices. These are technological and pedagogical packages. Mapping them respectively to active, constructive and interactive ICAP modes represents an intended affordance rather than observed cognition: the file does not contain the conceptual quality of wall posts, reciprocal exchanges or slide-viewing logs. The strongest incremental display contrast compares the slide and wall conditions, but that contrast also compares classes E/F with C/D and can capture unrecorded differences between them.

The revised file includes nine 1–5 responses for an adapted Engaged Learning Index: three each for cognitive, affective and behavioural engagement. We recomputed each domain as its three-item mean and the global score as the mean of those domains. The companion sheet provides Portuguese and English wording and the scoring rules. Because several items explicitly mention Wooclap, screen focus or avoiding off-task activity, the adaptation may respond differently across configurations even at equal underlying engagement. It should not be identified with the validated original instrument without evidence of translation, adaptation and measurement equivalence (Schreiner and Louis, 2011).

Four additional 1–5 items address fear of judgement, comparative exposure, inhibition of participation and disrupted focus; their unweighted mean is the comparison-threat score. The fourteen binary indicators record whether an off-topic activity category occurred, including messaging, news, games and unrelated generative AI use. We sum the indicators to obtain a breadth count from zero to fourteen. A yes/no field records voluntary participation; a 1–5 field records course interest; and another gives a mean assessment score. A count of categories cannot reveal switching frequency, duration of distraction or the educational relevance of a particular application use. No cognitive-load instrument or observed learning-process trace is present.

We calculated raw Cronbach's alpha from the item matrix using the variance of the item sum and the sum of the item variances. Alpha is a measure of internal consistency under assumptions, not a test of validity, unidimensionality or causal sensitivity (Tavakol and Dennick, 2011). The domain alphas and the four-item comparison-threat alpha are reported rather than copied from the earlier manuscript. The seven-sheet workbook is a substantive improvement in transparency because readers can inspect individual responses, but the source system, extraction date and link to the earlier workbook remain to be documented.

3.3 Analytical Approach

We report student-level means and standard deviations by configuration, Welch comparisons by feedback, one-way ANOVA and Tukey-adjusted pairwise contrasts as descriptive analyses. A two-way linear model assessed a configuration-by-seating interaction. For participation, we used the uncorrected Pearson chi-square calculation on the two-by-two table. All calculations were performed afresh from item responses, not inherited from displayed workbook formula caches. Two-sided p-values are shown where useful, but student-level p-values are not evidence that six allocated classes provide 303 independent experimental units.

For a cluster-aware sensitivity check, we averaged the outcome within each class and enumerated the 90 possible assignments of six classes into three labelled pairs of two. We compared the observed slide-minus-quiz difference with the two-sided distribution of all

such assignments. This exact test is valid as a randomisation test only if the asserted class-level randomisation really occurred with an exchangeable allocation set. The workbook does not document that process; accordingly, the result is conditional rather than confirmatory. For the incremental slide-versus-wall comparison, only four classes are involved and there are six possible pair allocations, making a conventional two-sided $p < .05$ impossible under that restricted exact design. We did not present a mediation test because exposure, proposed mediator and engagement are all post-only records without temporal ordering.

The estimand also changes when classes replace students as the unit of analysis. The student-level mean describes the average response among participants in the recorded classes. The equally weighted class-mean contrast describes the difference between average classes under the observed allocation. With approximately equal class sizes these estimates are numerically similar, but their uncertainty is very different. A cluster-level comparison does not manufacture information about what would have happened in unobserved classes. In particular, the observed feedback comparison cannot be made conditionally independent of class by adding class fixed effects, because feedback does not vary within any class. We report its percentages and score differences to characterise the dataset while withholding a feedback-specific causal estimand.

4. RESULTS

4.1 Data Checks and Psychometric Properties

The item matrices contain 303 complete records each. Recomputed individual domain scores and comparison-threat means follow the rules given in the new workbook. The revised values do not reproduce the previous extract by student ID: only 54 students have the same fourteen-item total across the two files, and the across-file correlation of those totals is approximately .07. The assessment fields likewise disagree for nearly every matched ID. These differences cannot be explained by spreadsheet display rounding. They require a provenance explanation before the authors designate one file as the definitive study export. The analyses that follow use only the revised file and retain version-specific denominators.

Internal consistency for all nine engagement items was $\alpha = .764$, but the three-item cognitive, affective and behavioural subscales yielded $\alpha = .475$, .631 and .490, respectively. Thus, the large domain-mean differences below should be interpreted with particular caution. The four comparison-threat items yielded $\alpha = .297$, far below the .85 claimed in the earlier article. Item means ranged from 2.22 to 2.31, yet their covariation was weak. A group difference in their average is a difference in an ad hoc composite, not a validated change in one coherent social-threat construct. Table 2 makes the distinction between available item responses and supported scale claims explicit.

Table 2. Item-level measures and internal consistency

Measure	Items	Scale	Recomputed alpha
ELI cognitive	3	1–5	.475
ELI affective	3	1–5	.631
ELI behavioural	3	1–5	.490
ELI all items	9	1–5	.764
Comparison threat	4	1–5	.297
Off-task categories	14 binary	0–14 total	Not assessed as a single latent scale

Alpha can vary with item count and dimensionality, and no single numerical threshold substitutes for a validity study. Nevertheless, the very low comparison-threat coefficient materially changes the evidential status of the feedback analysis. The item wording mixes personal apprehension, avoidance of speaking and loss of focus, which may be related yet non-identical responses. The source file allows item-specific and structural analyses; a new study should first examine the adapted wording, distributions and factor structure rather than using a previously asserted reliability value.

The distinction between numerical reproducibility and construct validity is central here. Recalculating a four-item mean from every row verifies arithmetic, while reliability asks whether the four responses behave coherently enough to justify summarising them as one variable. Validity further asks whether the resulting score represents the intended social-evaluative experience in this setting. The present records answer the first question, raise substantial concern about the second, and cannot answer the third from internal consistency alone. Because the group contrast is defined by a score with weak inter-item agreement, a large standardised test statistic may create a false impression of a precise psychological inference. Item-level reporting and an independently validated measure would be more informative than an additional decimal place in the group p value.

4.2 Configuration, Off-Task Activity and Engagement

The mean breadth of off-task categories was 4.71 (SD 1.61) for quizzes, 4.22 (SD 1.33) with the wall and 3.50 (SD 1.49) with slides. The student-level ANOVA was $F(2,300) = 17.24$, $p < .000001$. Tukey-adjusted differences were −1.22 for slides versus quizzes (95% simultaneous interval −1.71 to −.73), −.72 for slides versus the wall (−1.21 to −.23), and −.50 for the wall versus quizzes (−.99 to −.004). These are associations between configurations and reported category totals. In this revised file the incremental slide-versus-wall contrast is clearer than in the previous extract, which is one reason provenance matters.

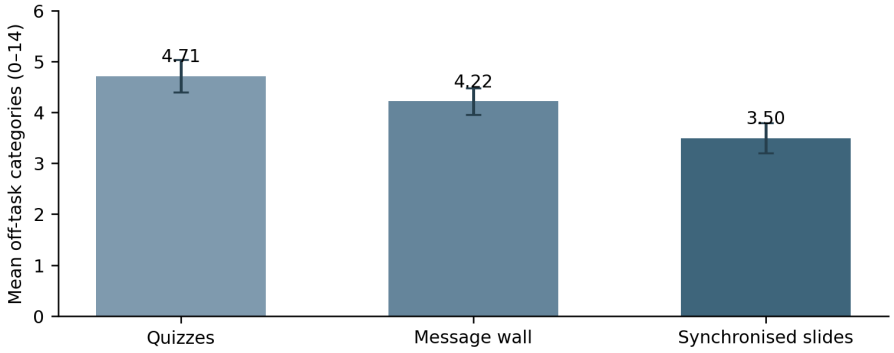

Figure 1. Mean off-task activity categories by classroom configuration; error bars are student-level 95% mean intervals, not cluster-adjusted uncertainty

The two class means were 4.84 and 4.58 in the quiz condition, 4.31 and 4.12 with the wall, and 3.45 and 3.54 with slides. The slide-minus-quiz difference in equally weighted class means was approximately −1.22. Under the hypothetical exchangeable 90-allocation randomisation set, its two-sided exact p value is 2/90 = .022. This conditional sensitivity result is compatible

with a class-level contrast but depends entirely on the unverified allocation claim. Restricted to the four wall/slide classes, the exact two-sided p value for the incremental contrast is 2/6 = .333. The latter does not erase the descriptive −.72 difference; it shows how little allocation-level information two classes per group can provide.

The global nine-item engagement mean rose from 3.34 (SD .24) to 3.69 (.19) and 4.22 (.21) across configurations. The student-level omnibus $F(2,300) = 445.40$, $p < .0001$. Cognitive means were 3.48, 3.81 and 4.29; affective means 3.17, 3.76 and 4.27; behavioural means 3.37, 3.50 and 4.10. The item-level global score is associated with condition, but domain reliability and condition-specific wording limit comparisons of the underlying constructs. Mean assessment scores of 3.62, 3.73 and 3.77 were much closer, with $F(2,300) = 1.18$, $p = .309$. Table 3 and Figure 2 place reported engagement beside achievement to avoid equating the two.

Table 3. Revised workbook outcomes by configuration (101 students per condition)

Outcome	Quizzes M (SD)	Wall M (SD)	Slides M (SD)
Off-task categories	4.71 (1.61)	4.22 (1.33)	3.50 (1.49)
ELI cognitive	3.48 (.39)	3.81 (.31)	4.29 (.33)
ELI affective	3.17 (.37)	3.76 (.33)	4.27 (.33)
ELI behavioural	3.37 (.42)	3.50 (.35)	4.10 (.35)
ELI global	3.34 (.24)	3.69 (.19)	4.22 (.21)
Assessment (1–5)	3.62 (.75)	3.73 (.75)	3.77 (.60)

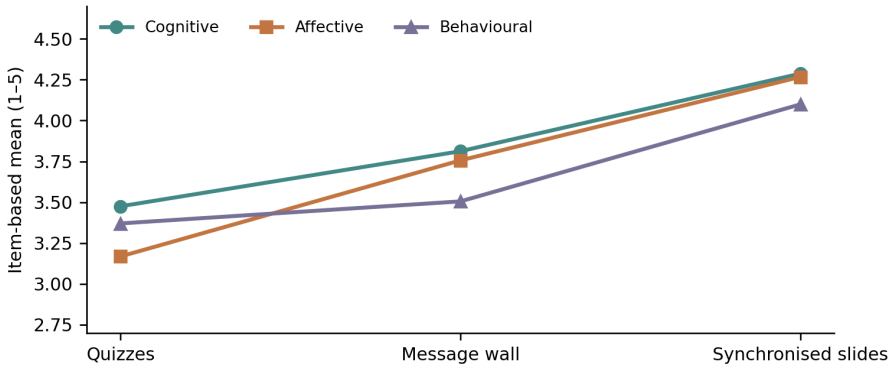

Figure 2. Mean item-based engagement scores across the three cumulative configurations

Across students, the global engagement score correlated only weakly with assessment ($r = .074$), and the off-task total with assessment at $r = -.032$. These correlations do not identify whether involvement improves knowledge; assessment may reflect prior attainment, different tasks or measurement error. They do challenge an assertion that the large differences in self-reported engagement were accompanied by comparably large differences in measured performance. The immediate pedagogical implication is to distinguish learners' perceived participation from demonstrable mastery.

4.3 Seating and the Proposed Spatial Pathway

Seating was recorded as front, middle or back. The condition-by-seat interaction for the off-task total was $F(4,294) = 1.57$, $p = .182$. In the slide condition, front, middle and back means were 3.94 ($n = 31$), 3.15 ($n = 41$) and 3.52 ($n = 29$). In the quiz condition, the corresponding means

were 4.56 (n = 34), 4.97 (n = 38) and 4.55 (n = 29). The revised values differ materially from the earlier manuscript's figure, and none establishes equivalence between seating locations. The small number of people in each cell and non-significant interaction leave the accessibility account unconfirmed.

A back-row contrast is not a direct measure of projector visibility. Seat choice may reflect preference or constraints, and there is no individual record of whether a student used the synchronised display. A pre-specified equivalence margin, confidence interval and measured visibility would be needed to claim spatial parity. The corrected descriptive pattern is displayed in Figure 3 without a "parity achieved" label. It may motivate an accessibility-focused experiment, but it cannot establish that an interface neutralises all seating disadvantages.

4.4 Feedback, Participation and Assessment

Collective feedback appeared in 151 records and private feedback in 152. Comparison-threat composite means were 2.50 (SD .36) and 2.01 (.34), respectively; a student-level Welch test gave $p < .0001$. Interest was 4.74 (.28) versus 4.56 (.36), $p < .00001$. Voluntary participation occurred in 36/151 (23.8%) versus 58/152 (38.2%), uncorrected Pearson $\chi^2(1) = 7.26$, $p = .0071$. Assessment means were 3.59 (.71) versus 3.83 (.68), $p = .0028$. These calculations are reproducible from the revised rows and supersede the earlier extract's feedback percentages and scale means.

Table 4. Feedback patterns in the revised workbook

Outcome	Private (n = 152)	Collective (n = 151)	Descriptive contrast
Comparison threat	2.01 (.34)	2.50 (.36)	+.49; low reliability
Course interest	4.56 (.36)	4.74 (.28)	+.18
Voluntary participation	58 (38.2%)	36 (23.8%)	−14.3 percentage points
Assessment	3.83 (.68)	3.59 (.71)	−.24

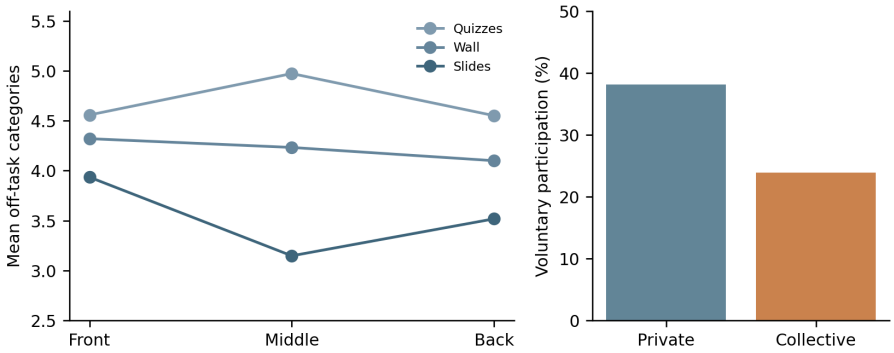

Figure 3. Off-task categories by seat and configuration;
separate panel shows voluntary participation by feedback class

The feedback comparison has two simultaneous weaknesses. First, the four-item comparison scale has $\alpha = .297$, so the numerical difference is not strong evidence about a unified anxiety construct. Second, feedback modality is perfectly aligned with the six class labels: in each configuration one entire class is private and the other collective. The apparent differences may

reflect class composition, instructor or time slot. The sample contains only three classes per feedback modality and one of each within a configuration. Treating the 303 students as independently assigned to feedback is pseudoreplication. These patterns are informative hypotheses about response visibility, not identified effects of public display.

The gender distribution of 226 male, 68 female and nine other/prefer-not-to-say students offers no basis for a confirmatory gender interaction, especially with only four or five students in the smallest feedback-by-gender cells. We do not interpret those cell means. Future comparisons should independently manipulate visibility within classes or recruit substantially more classes, retain item-level data, and predefine whether the main outcome is willingness to speak, the quality of explanations, or validated social-evaluative concern.

5. DISCUSSION

5.1 What the Revised Evidence Supports

The revised dataset exhibits a coherent descriptive gradient in off-task categories and item-based engagement across cumulative digital classroom configurations. Both classes in the slide condition had lower off-task means than both classes with quizzes alone, and the condition contrast remains notable in a conditional class-level randomisation enumeration. At the same time, the incremental comparison with the wall is based on only two classes on either side, so its small student-level p value overstates the certainty supplied by the actual allocation units. The present evidence can motivate more targeted trials; it cannot isolate the slide interface from class effects or prove that making the learning screen salient raised switching costs.

The distinction between the two Excel versions is a scientific issue, not a formatting discrepancy. Identical IDs do not denote identical outcomes or assignments across them. A journal submission must identify the definitive extract, account for any transformation of source records, and preserve a reproducible audit trail. The revised workbook has item-level data and class-consistent assignments, but that fact alone does not authenticate their origin. A documented export from the survey and response platform, allocation list, raw attendance record and dated codebook would allow an editor to assess what was observed. Until then, terms such as “experimentally demonstrated”, “replication” and “verified randomisation” would be premature.

The new item-level access is valuable because it exposes limitations that earlier summary statistics concealed. The global engagement index has moderate consistency, but its component domains have weak-to-moderate alpha values, and the proposed comparison-threat inventory has low consistency. An adapted item that directly asks whether a student stays focused on the screen may partly measure the intervention’s presentation rather than a general behavioural trait. A psychometric follow-up should examine wording, translation, item-total relationships, dimensionality and invariance across configurations. Scale development may be needed before using these scores to assess differential socio-emotional responses.

5.2 Pedagogical Design without Mechanistic Overclaim

The practical case for displaying slides on a personal device is strongest where the projection is difficult to see, students need to inspect a detailed diagram, or the interface supports a prompt that demands explanation. These design benefits are plausible regardless of the study's unresolved causal status. A future factorial comparison could keep quizzes, wall tasks, instructor, content and feedback stable while varying slide synchronisation. Direct records of device switching, task completion and visual accessibility could then separate the hypothesised "interface occupation" mechanism from better access to content. Measuring short-term knowledge and delayed transfer would clarify whether reported engagement reflects learning gains.

The ICAP framework sharpens this design. A quiz click can demonstrate active manipulation; a justified answer may become constructive; peer discussion becomes interactive only when learners respond to and build on one another's ideas (Chi and Wylie, 2014). The current dataset counts none of these processes. Future studies should code a sample of message-wall posts and discussion transcripts, record when students actually view slides and relate these processes to assessment aligned with the four lesson topics. A platform feature is an opportunity for a cognitive mode, not evidence that the mode occurred.

The feedback patterns raise a separate question about psychological safety. An anonymous class distribution can expose a useful conceptual disagreement without ranking individuals, whereas a public performance comparison may discourage uncertain learners from speaking. Instructors can test both formats in low-stakes exercises, ask for written explanations before public discussion, and offer private routes to challenge a displayed result. These recommendations arise from pedagogical reasoning and the descriptive trade-off, not from proof that feedback visibility caused the observed score differences. The weak threat scale and class confounding make careful wording essential.

Accessibility should also be separated from attention control in a future implementation. A student may benefit from enlarged code, diagrams or captions on a personal display while continuing to receive unrelated notifications. Another may use a phone whose small screen makes detailed slides less accessible than the wall projection. A design study could measure device type, screen size, visibility, notification practices and actual engagement with each slide, then estimate whether the same interface benefits all learners. If content access improves without a reduction in switching, that would still be a meaningful educational outcome. The mechanism should follow the evidence, rather than requiring all benefits to be interpreted as prevention of multitasking.

5.3 Inference, Ethics and the Next Study

A definitive evaluation needs more allocation units and a documented assignment procedure. If classes are randomised, the reporting should include eligibility, recruitment, class and student flow, allocation generation, consent timing, baseline characteristics by cluster, intraclass correlation assumptions and cluster-aware analysis, as recommended for cluster trials (Campbell et al., 2012). A stepped or counterbalanced arrangement across sessions could reduce the dependence of the configuration effect on persistent class differences, provided carryover and topic effects are planned for. Under either design, analysis should designate a primary outcome and a minimally meaningful difference before observing the results.

The supplied workbook asserts that a committee approved the study under reference CEIHD-2025/084 on 15 September 2025. It provides no signed decision or institutional correspondence, and the committee is identified only generically as belonging to a “university institution”. This is a claim within a spreadsheet, not independently verified evidence of approval. The actual approving body, formal reference, protocol scope and date must be checked against the original institutional decision. Likewise, consent to participate does not automatically authorise public release of student-level responses. Any repository deposit should apply the approved consent terms, disclosure-risk assessment and access controls.

These documentation needs do not negate the value of the observed records. They define the boundary between an internally reproducible analysis and a publishable report that an external reader can trust. The revised file enables independent recomputation of item-based scores and highlights the methodological decisions that matter most. Its next version should pair a traceable export with a reliable scale, sufficient clusters and transparent ethical approval. That combination would transform a promising classroom pattern into a persuasive contribution to digital pedagogy.

6. CONCLUSION

Attention in a digital lecture is shaped by the work that appears on students’ devices, the access those devices give them to course material, and the social meaning of a visible response. In the revised records, classes using quizzes, a message wall and synchronised slides reported fewer categories of off-task activity and higher item-based engagement than classes using quizzes alone. The pattern is consistent across the two classes at each extreme and supplies a concrete reason to study the design of the student-facing screen. Assessment differences were small, however, and the six-class design cannot establish that the slides themselves caused the observed gradient.

The strongest contribution of the item-level workbook is the precision it brings to this interpretation. It reveals modest internal consistency for the engagement subscales and poor consistency for the comparison-threat inventory. It also shows that collective feedback is aligned with class membership, so its links with interest, participation and assessment cannot be attributed to visibility alone. The seating interaction gives no secure basis for a spatial-equality claim. These results refine the research questions: does personal display improve access or reduce switching, what kinds of prompts elicit constructive and interactive reasoning, and under which conditions does feedback visibility encourage discussion without discouraging participation?

A subsequent study can answer those questions by varying interface features independently, documenting class allocation, retaining time-resolved process measures, validating the adapted instruments and evaluating learning beyond immediate self-report. For current practice, the evidence supports careful design and evaluation of screen content, generative prompts and response visibility. For publication, the revised extract must be reconciled with the earlier file and the institutional approval verified against the original decision. That transparent chain from source data to interpretation will make the findings more useful to instructors and more credible to readers.

ETHICS AND DATA AVAILABILITY

The revised workbook reports ethics reference CEIHD-2025/084 and a decision date of 15 September 2025. The workbook contains pseudonymous item-level data for the nine engagements, fourteen off-task and four comparison-threat indicators. Public sharing of the individual records requires confirmation that the consent and approval permit it, alongside a disclosure-risk assessment.

ACKNOWLEDGEMENTS

This paper is a revised and expanded version of a paper entitled “AN ICAP-BASED APPROACH TO REDUCING MULTITASKING IN DIGITAL CLASSROOMS” presented at ELDL2026, Valencia, Spain, 25 to 27 July 2026.
The authors have used AI technology for grammar checking and for managing references. They used Grammarly and NotebookLM Gemini. Participants provided informed consent on a Microsoft Forms questionnaire, and the result file is kept by the authors. The participants' data has been anonymised and the article does not include any potential identifiers.
All authors declare that they have no conflicts of interest.

REFERENCES

Campbell, M. K., Piaggio, G., Elbourne, D. R. and Altman, D. G. (2012). Consort 2010 statement: extension to cluster randomised trials. *BMJ, 345*, e5661. doi: 10.1136/bmj.e5661

Chi, M. T. H. and Wylie, R. (2014). The ICAP framework: linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243. doi: 10.1080/00461520.2014.965823

Fredricks, J. A., Blumenfeld, P. C. and Paris, A. H. (2004). School engagement: potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. doi: 10.3102/00346543074001059

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H. and Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23). 8410–8415. doi: 10.1073/pnas.1319030111

Hutain, J. and Michinov, N. (2022). Improving student engagement during in-person classes by using functionalities of a digital learning environment. *Computers & Education, 183*, 104496. doi: 10.1016/j.compedu.2022.104496

Hutain, J. and Michinov, N. (2024). Collective feedback based on quizzes in online learning: a “double-edged sword” effect on attitude to courses, emotions, and academic behaviors. *Interactive Learning Environments, 32*(5), 1894–1906. doi: 10.1080/10494820.2022.2131831

Jamet, E., Gonthier, C., Cojean, S., Colliot, T. and Erhel, S. (2020). Does multitasking in the classroom affect learning outcomes? A naturalistic study. *Computers in Human Behavior, 106*, 106264. doi: 10.1016/j.chb.2020.106264

Kahu, E. R. and Nelson, K. (2018). Student engagement in the educational interface: understanding the mechanisms of student success. *Higher Education Research & Development, 37*(1), 58–71. doi: 10.1080/07294360.2017.1344197

Michinov, N. and Hutain, J. (2024). Displaying the teacher's slideshow on students' devices prevents multitasking and promotes engagement during lectures. *Active Learning in Higher Education, 25*(3), 551–567. doi: 10.1177/14697874231176987

Sana, F., Weston, T. and Cepeda, N. J. (2013) 'Laptop multitasking hinders classroom learning for both users and nearby peers. *Computers & Education, 62*, 24–31. doi: 10.1016/j.compedu.2012.10.003

Schreiner, L. A. and Louis, M. C. (2011) 'The Engaged Learning Index: implications for faculty development. *Journal on Excellence in College Teaching, 22*(1), 5–28.

Sweller, J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science, 12*(2), 257–285. doi: 10.1207/s15516709cog1202_4

Tavakol, M. and Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. doi: 10.5116/ijme.4dfb.8dfd