

Tiers of Engagement III: The Psychological Learning Engine in Online Business Simulators

Martin Guttormsen¹ and Jørn Weines²

¹School of Economics and Business, UiT The Arctic University of Norway, Alta, Norway

²Centre for Teaching, Learning and Technology, UiT The Arctic University of Norway, Tromsø, Norway

martin.guttormsen@uit.no

jorn.weines@uit.no

Abstract: This article explores the primary drivers of student learning in the context of business simulation games as part of an online business economics program. Specifically, it investigates whether perceived learning is driven more by a student's internal psychological states or by external course designs, such as added narrative assignments and optional collaboration. By analysing quantitative and qualitative survey data (N=321) in the form of a post-simulator reflection assignment, the study evaluates the impact of motivation and engagement on business simulators as learning environments. The findings reveal that students' internal psychological states, specifically their intrinsic motivation, active engagement, and depth of reflection make up the primary "engine of learning". Demographic factors, such as gender and age, had no significant impact. Pedagogical elements external to the simulator, like narrative assignments and optional collaboration, did not directly improve perceived learning. However, they functioned as essential catalysts. Collaboration acted as a "social buffer" that helped students overcome technical and academic frustrations. Added narrative tasks acted as "metacognitive speed bumps" that gave motivated students opportunities to switch between types of learning activities, enabling them to connect the simulator to the theoretical curriculum. On the other hand, students with low motivation experienced the tasks as disruptive. The study concludes that while intrinsic motivation and engagement drive learning, thoughtfully designed friction can help in fostering deep, self-regulated learning in online simulators.

Keywords: Game-based learning, Business simulators, Self-determination theory, Metamotivation, Self-regulated learning, Online education

1. Introduction and Theoretical Background

This article is our capstone study on the use of business simulators in a fully online economics program. In our previous articles (Guttormsen and Weines 2024, Guttormsen and Weines, 2025), we presented and refined our "Tiers of Engagement"-framework. This approach groups students based on their reported presence of intrinsic motivation and engagement. We use these tiers to examine differences in the depth of reflection on learning, reported perceived learning, and utility of added narrative tasks across tiers. Our framework was inspired by Schwartz and Plass' (2019) work on types of engagement in learning with games, Deci and Ryan's (2000) self-determination theory (SDT) and Miele and Scholer's (2018) concept of metamotivational processes regarding students' self-regulation of learning (SRL).

Our context of study is a flexible, fully online bachelor program in business economics and administration. In the course on business economics analysis, the "Hubro Business Simulator" is used as a learning tool. In this simulation game, the students manage start-up companies that invest in the development, production and sales of jetpacks. Over the course of the simulation, students make decisions on investment in research, production capacity and marketing. In addition, the students get narrative tasks framed as assignments from the board of directors, where they need to perform business economics analysis.

This article is based on a survey of student responses (N=321), with both quantitative and qualitative questions. We also draw on our two previous datasets (N=328 and 388) for comparison. The main purpose is to validate our framework for categorizing student effort and motivation towards simulators as a learning activity. To replicate the framework and descriptive analysis from earlier studies we use the same criteria as before (Appendix 1). To validate our framework, we performed a statistical analysis where we examine significant differences between student groups and significant correlations between measured variables and factors.

In our previous analysis we found a pattern of declining reported perceived learning and reflection, which illustrated that high engagement and intrinsic motivation both increased reported achievement of learning objectives, reflection on learning and the utility of the narrative tasks. In the analysis of the qualitative data, we found that the students with the highest engagement and intrinsic motivation exhibited clearer views of the simulator as a meaningful and valuable learning activity. We refer to our aforementioned works for further details (Guttormsen & Weines, 2025).

To elaborate our findings, we draw on students' qualitative comments, aiming to paint a more nuanced picture of how students experience online business simulators as learning arenas. Our goal is that this can serve as a bridge between our quantitative and qualitative data, as well as increase the transferability of our framework to other approaches in game-based learning.

In this article, we further explore how the simulation game fits in with the overall learning design of the course. To do this, we rely on the Conversational Framework (Laurillard, 2012). This framework builds on the collected literature of students' learning to represent the different roles teachers and learners play, to visualize "what it takes to learn, and therefore what it takes to teach" (ibid., p. 93). In addition to serving as a tool for design, the framework can be applied to analyse the combination of learning types an activity combines and motivates cycles of. The framework aims to visualize how these cycles in a teacher's learning design motivate or enable learners to generate articulations and activity, that in turn help shape their concepts and practice (ibid., p. 94). The use of the simulator in our study has been intentionally designed as a learning experience that fits the online nature of the course, where the students have a long period of time to complete the simulator itself and the accompanying narrative tasks. An explicit goal of this is to provide simulated practice that allows students to bridge theoretical concepts with practical exercises. In analysing and discussing our findings we will draw on the conversational framework to further explore how students experience the simulator, peeking behind the quantitative metrics. Therefore, our primary research question is:

In the context of an asynchronous, online business simulator, how do students' internal psychological states (intrinsic motivation and engagement) interact with the pedagogical design (narrative tasks and optional collaboration) in driving self-reported perceived learning and metacognitive reflection?

2. Methodological Design

2.1 Research Design and Sample

This study employs a concurrent triangulation design (Almeida, 2018), a mixed-methods approach where quantitative and qualitative data are collected concurrently and analysed to validate, complement, and elaborate on one another. The sample consists of students (N=321) enrolled in an online bachelor's program in business administration and economics who completed the "Hubro Business Simulator" as a mandatory coursework assignment. The sample is characterized by an almost perfect gender balance, comprising 157 men (48.9%) and 164 women (51.1%). A majority of the respondents are mature students with an average age of approximately 29 years, which is typical for online and flexible degree programs.

We screened the dataset for missing data prior to analysis, revealing minimal missingness across variables (< 5%). Two responses (0.6%) for the 'Collaboration' variable were excluded due to ambiguous, uncodable data. For 'Degree of intrinsic motivation', five respondents (1.5%) selected 'none of the above,' rendering the scale structurally non-applicable. Finally, for the 'Narrative assignments' variable, 15 responses (4.7%) were excluded because the qualitative data indicated a fundamental misunderstanding of the question prompt, rendering the data invalid. As the missing data was consistently under the 5% threshold and driven by structurally missing, uncodable or invalid responses, imputation methods were deemed inappropriate. Instead, available-case analysis (pairwise deletion) was utilized to maximize statistical power, resulting in minor fluctuations in sample sizes across specific analyses (ranging from N=306 to N=321).

2.2 Measurement Instrument

The data are based on a post-simulation survey, framed as a reflection assignment. The instrument consisted of quantitative questions using 5-point Likert scales, alongside open-ended free-text questions. Self-reported data are considered valid, provided the questions are unambiguous and relate to the respondent's recent, direct experiences (Gonyea, 2005). The questions in the survey are reported in Table 1 below.

Table 1: Survey question and coding scale

Questions	Possible answers	Coded data for analysis
Participation in research	I agree / I do not Agree	1, 2
Gender	Male, Female, Other	1, 2, 3
Age	Scale (18 – 100)	Number 18 - 100
Effort put into the simulator (proxy for engagement)	Scale (1 – 5)	1-5
Experience with the simulator (proxy for intrinsic motivation)	Very boring/uninteresting(1), a bit boring/uninteresting(2), neither boring/uninteresting or fun/interesting(3), a bit fun/interesting(4), very fun/interesting(5), none of the above(6)	1-5 (respondents answering 6 was filtered out due to uncodable non-applicable data).
Reflections regarding using the simulator	Free text	1-3
Discussion/collaboration with other students and learning	Free text	1-3
Practical relevance and feedback	Free text	Used as assistance in coding reflection
Achievement of learning goal “Knowledge of the role of economics and function in a business and ability to discuss foundational elements in economic business management”	Scale (1-5)	1-5
Achievement of learning goal” Knowledge of the cost term and understanding of cost variability”	Scale (1-5)	1-5
Achievement of learning goal “Be able to conducts simple analyses of optimal product choices in situations with limited resource access”	Scale (1-5)	1-5
Achievement of learning goal “Have knowledge of different perspectives on price setting in practice, in different market forms”	Scale (1-5)	1-5
Narrative assignments and learning	Free Text	1-3
Other feedback	Free text	Used as assistance in coding reflection

2.3 Quantitative Analysis Method

2.3.1 Descriptive analysis in Excel

Following data collection, several raw variables necessitated recoding to establish appropriate parameters for the empirical analysis. We describe this process in section 2.4., and the specific criteria applied for this categorization are reported in Appendix 1.

The tier-based analytical framework was developed by combining two primary self-reported metrics: the students' perceived intrinsic motivation towards the simulator, and their level of effort (serving as a proxy for engagement). Intersecting these two dimensions generated a comprehensive matrix of 15 distinct engagement tiers based on the specific combination of motivation (High, Medium, Low) and engagement score (1-5). For example, students exhibiting a high presence of intrinsic motivation and the highest level of effort are grouped as tier **H5**, while students with low intrinsic motivation and the middle level of effort are grouped as **L3**. To determine the population distribution, each tier was subsequently calculated as a proportion of the overall sample size.

These tier classifications also served as the primary grouping variables for the subsequent descriptive analysis. Segmenting the data in this manner enabled the computation of descriptive statistics, namely means, standard deviations, and relative frequencies for each specific group. The variables assessed across these tiers included the students' depth of reflection, perceived attainment of learning outcomes, demographic characteristics (age

and gender), and the perceived educational value derived from the narrative assignments. These metrics were systematically examined to identify broader underlying patterns.

2.3.2 Statistical analysis in Jamovi and technical details

All statistical analyses were conducted using the free, open-source statistical software platform Jamovi (The jamovi project, 2025). We used Jamovi because it has an easy-to-use interface like SPSS, but as it utilizes R coding, it provides flexibility and accuracy.

Prior to conducting the inferential statistics, specific data transformations were applied in Jamovi. To meet the strict variance assumptions required for parametric testing (Welch's One-Way ANOVA), the two lowest categories of the independent variables Engagement and Motivation were collapsed due to extremely low response rates that resulted in the ANOVA being unable to run. However, the original uncollapsed 5-point scales were retained for correlation analyses to maximize data variance.

To validate that the four questions relating to self-perceived achievement of the learning outcomes measured the same underlying construct (perceived learning from the simulator), we conducted an exploratory factor analysis (EFA). Principal Axis Factoring (PAF) was chosen as the extraction method. PAF is recommended over Principal Component Analysis (PCA) in social science research when the objective is to identify latent psychological variables (Costello and Osborne, 2005; Fabrigar et al., 1999). Since latent factors in learning environments are rarely entirely uncorrelated (Costello and Osborne, 2005), Oblique rotations are deemed most appropriate for behavioral data. We therefore used Oblimin in our analysis.

To analyse group differences, we employed Welch's One-Way ANOVA to adjust for unequal variances and protect against Type 1 errors (Delacre et al., 2019). We ran ANOVAs between each independent variable and the factor/dependent variable. This allowed us to examine if different responses on the independent variables had an impact on perceived learning from the simulator. The rationale for deviating from the classic Fisher's ANOVA was the presence of highly unequal group sizes and variances in the dataset. The significance level was set at $\alpha=.05$. Finally, Spearman's rho (ρ) was used for correlation matrix analyses, as it is a robust non-parametric test suitable for ordinal Likert data (Field, 2018).

2.4 Qualitative Analysis Method

As described in more depth in our previous articles, the analysis method and framework are anchored in SDT (Deci and Ryan, 2000), SRL and metamotivation (Miele and Scholer, 2018). The open-ended free-text responses (N=1605 text comments) were manually analysed by the authors working as a team, using the criteria reported in Appendix 1 to code the depth of the students' metacognitive reflection, collaboration and perceived utility from the narrative assignments. Throughout the coding process, we made notes of observations from the data in our working document. A supplemental analysis assisted by a Large Language Model (Google Gemini Pro) was done to aid in pattern recognition, summary of reported user experiences with the simulator, and in constructing and describing the characteristics of each tier. All findings generated with assistance were manually verified against the raw data.

3. Empirical Findings and Analysis

Our EFA analysis extracted one distinct factor that explained 69.2% of the variance for the four questions related to achievement of specific learning goals. This extracted factor "Perceived learning from the simulator" demonstrated very high internal consistency, confirming our instrument's reliability for subsequent analyses.

The analyses of variance (ANOVA) reveal that students' internal psychological states (intrinsic motivation and engagement) are strong, statistically significant drivers of perceived learning, whereas external and demographic variables are not. We report the results of our analysis, showing the F-value (F), degrees of freedom between categories (df1) and within categories (df2). For engagement, a post-hoc Games-Howell test (Games and Howell, 1976) revealed that categories 4 and 5 were significantly different from each other ($p<.001$), indicating that incremental steps of high engagement yield measurable benefits.

The analysis shows that both intrinsic motivation and engagement levels have significant effects on reported perceived learning. Students with the highest levels in these variables also express the highest perceived learning. Similarly, we find that the depth of reflection on learning also has a significant impact on perceived learning. However, here we find a non-linear trend: While the lowest reflection level yielded the lowest perceived learning, students at the medium reflection level reported marginally higher perceived learning than students with the highest reflection level.

Our analysis also showed some non-significant results. The reported perceived learning was entirely independent of gender. Men and women performed almost identically. Furthermore, neither collaboration nor the perceived utility of the narrative assignments yielded direct significant effects on perceived learning from the simulator. While these results are not significant, they might point to some links to the other variables.

The results of our analyses are reported in Table 2 below.

Table 2: Results of statistical analysis

Category	Key Metrics & Statistics
Sample & Demographics	N=321 (Men: 157/48.9%, Women: 164/51.1%). Age: ~29 Comments analysed: 1605.
Data Cleaning & Validation	Missing data <5% (Final N: 306–321). EFA factor explained variance: 69.2%. Factor loadings: Ranging from 0.798 to 0.864 KMO=0.844, $p < 0.001$. Bartlett's Test of Sphericity: $\chi^2=775$, $df=6$, $p < .001$ Cronbach's $\alpha=0.899$.
Significant drivers (ANOVA) for average learning goals	Motivation: $F(3,39.8)=38.43$, $p < 0.001$ (High M=4.23, Low M=2.57). Engagement: $F(3,43.7)=20.03$, $p < 0.001$ (High M=4.35, Low M=2.89). Reflection: $F(2,123)=17.62$, $p < 0.001$ (Low M=3.65, Med M=4.16, High M=4.10).
Non-Significant drivers (ANOVA) for average learning goals	Gender: $F(1,318)=0.330$, $p=0.566$ (Men M=3.93, Women M=3.88) Collaboration: $p=0.098$, Narrative assignments: $p=0.98$
Correlations (Spearman)	Collaboration: w/ Engagement ($\rho=0.172$, $p=.002$), Motivation ($\rho=0.119$, $p=.035$), Reflection ($\rho=0.157$, $p=.005$) Narrative assignments: w/ Reflection ($\rho=0.345$, $p < 0.001$)

3.1 The Underlying Behavioural Network

By looking at the correlations between the quantitative data and triangulating with findings from the qualitative data, we can shed light on the underlying behavioural network of our respondents.

Spearman's rho and qualitative comments explain why external variables (such as collaboration and narrative assignments) lack a direct effect. Although “collaboration” did not drive the score directly, the correlation matrix shows that it functions as a catalyst that correlates positively with engagement, motivation, and reflection. Our analysis of the qualitative data supports this; the subset of students who report completing the simulator alongside others (8%) reported valuable academic discussions that helped resolve technical problems, academic frustration and exploring other perspectives.

One of our most central findings concerns the narrative assignments. Descriptive qualitative data indicates strong polarization: While unengaged students experienced the added tasks as “disruptive” interruptions, students in higher tiers claimed the tasks were essential for their learning. The correlation matrix explains this, showing a strong, statistically significant correlation between the narrative assignments and Reflection level. The assignments do not generate learning directly, but for some of the students they seem to prompt metacognitive reflection processes, such as (re)engagement with other learning material (lecture notes, recorded lectures or textbooks).

Qualitatively, students generally reported a high perceived realism in the economic mechanisms (supply and demand, effects of pricing), despite the product (“jetpacks”) being fictitious. As in our previous studies, many report positively on being able to experience the direct consequences of their actions. Concurrently, the analysis revealed two clear sources of noise and friction. First, linguistic and technical barriers (English interface, punctuation issues with dots/commas). Second, a tendency towards engaging with the simulator more as a game than a learning activity, with some students clearly stating that their primary focus was optimizing the simulator's numbers (i.e. succeeding in the game) rather than acquiring deeper academic understanding.

4. Limitations and Discussion

4.1 Limitations

Our study has several limitations that must be acknowledged. First, the findings are based on self-reported data, which are susceptible to systemic reporting bias (Gonyea, 2005). The high internal consistency (Cronbach's $\alpha=0.899$) mathematically indicates validity, but the qualitative comments reveal that some students engage in superficial "gaming of the system" to achieve fictitious profit. There is a risk that some students confuse winning the game with actual academic competence building. Second, the data show that linguistic barriers (English interface) and inflexible technical input requirements (dots versus commas) created cognitive overload for a portion of the sample. Such technical noise may have caused some students to end up in a lower motivation category in our analysis than they strictly belong to academically. Finally, to meet parametric assumptions for ANOVA testing, we collapsed the lowest response categories for engagement and motivation. While methodologically sound, this slightly reduces the granularity of the data at the lower extremes. Additionally, the use of available-case analysis (pairwise deletion) to handle structurally missing data resulted in a fluctuating sample size across tests (ranging from $N=306$ to $N=321$). While this approach successfully maximized statistical power, it means the specific subsamples utilized for different analyses are not perfectly identical.

As noted in section 2.3.1., the categories "Depth of reflection on learning", "Collaboration" and "Utility from narrative tasks" are not direct measurements or student reported but based on our analysis. When coding, we have worked as a team, following the criteria reported in Appendix 1, but we acknowledge the possibility of human error in the manual coding process. Due to the demographic profile of our students (high mean age, many combining employment and studies) and the fully flexible and online nature of the course, our findings might not be universally transferable to other contexts.

4.2 Discussion

This study has investigated the mechanisms driving students' learning when using business simulators in an asynchronous, online educational context. Through the triangulation of quantitative and qualitative data, a complex picture emerges wherein students' internal psychological states constitute the engine of learning, while pedagogical frameworks act as indirect facilitators.

In our previous study, we found a declining wave pattern, showing that the reported achievement of learning outcomes, depth of reflection and utility of narrative assignments were more driven by presence of intrinsic motivation than reported effort. In the current study, we also see a similar declining wave pattern for perceived learning, but not for depth of reflection and utility from narrative assignments.

As part of our qualitative analysis, we profiled the characteristics of the different tiers. The resulting characteristics were similar to our previous study (Guttormsen & Weines, 2025). In both studies, the top tiers represented about 50% of the sample. They are deeply immersed, use the simulator to practice strategic thinking, and successfully bridge practice with theory. The middle tiers do not report the same immersion but still tie theory to practice and express that they see the simulator as relevant. The lower tiers seem to be driven purely by compliance (passing the course). They experience cognitive overload, complain about the English interface (having not found the setting for language), and find the simulator repetitive and frustrating.

We find the same polarisation towards the narrative assignments as earlier, where high tiers experience them as bridges between theory and practice, while low tiers see them as technical friction. This also ties in with the statistical analysis and the correlation between the narrative assignments and reflection.

However, some differences from the earlier study emerges as we shed new light on the mechanisms in play, especially from the statistical analysis. For instance, the role of gender and age, which we theorized could be influencing performance, seem to completely disappear in the data in this study. This is also backed by these variables having no significant differences between the groups in the ANOVA, despite a slight negative correlation between women and intrinsic motivation in the correlation matrix.

Our current analysis finds that intrinsic motivation serves as the strongest driver. Our findings for both engagement and motivation provide solid empirical support for earlier assumptions within the field of Game-Based Learning, i.e. that active engagement and intrinsic motivation are prerequisites for deep learning (Schwartz and Plass, 2019). This aligns strongly with the SDT framework (Ryan and Deci, 2000). Unlike other studies (Butler et al., 2025), we do not find that men benefit more from simulator-based education than women. We find a small but significant difference, with men reporting a slightly higher intrinsic motivation, but despite this both genders report identical perceived learning. With a sample size of 321 students featuring perfect

gender balance, our data establishes that the simulator functions as a gender-neutral arena. Our statistical analysis shows that motivation and effort, not demographic factors, drive the outcomes. It is possible that the mature students in our context, a fully online, flexible program have a different motivational profile that overrides the effect of gender.

As mentioned in the introduction, we draw on the Conversational Framework for analysing how students experience the business simulator and its place in the overall learning design of the course. This framework highlights six broad groups of learning types: Acquisition, inquiry, practice, production, discussion and collaboration (Laurillard 2012, p. 96). As a fully flexible online course, the overall learning design focuses heavily on individual learning through acquisition (learning materials and video lectures via the learning management system) and practice (through assignments and the simulator). There are optional synchronous webinars (combining practice and discussion), where students can interact with the instructor. These are also available as recordings for later (re)viewing. Self-organized activities (i.e. discussion and collaboration) between students are encouraged, but as discussed in our previous articles, the students have chosen to attend the program due to their need for a high degree of flexibility (for instance due to combining studies with full-time employment).

The qualitative data leads us to an interesting observation regarding collaboration. Over the course of our three studies, many students report that while discussion and collaboration could be beneficial, they see it as practically unfeasible. In the current study, 15% report having had some degree of discussion or collaboration in connection with the simulator. A slight majority of these also clearly state that they perceive a positive impact on their learning from having done so. Some of those who indicate no collaboration express that they still think it would have been beneficial. In our statistical analysis, we find no direct impact from collaboration on perceived learning, but the correlation matrix clearly shows that it serves as a positive catalyst for motivation, engagement and reflection.

The simulator activity is not intended as a single standalone opportunity for practice. The instructor has made an additional layer of narrative tasks, framed as assignments from the company's board, intended to provide opportunities to engage with the course learning material while progressing through the simulator. The intent is assisting in bridging simulator practice with the theoretical curriculum. In our analysis, we found no direct statistical correlation on students' perceived learning from these assignments. However, the quantitative correlation with reflection, supported by the qualitative analysis, demonstrates that these tasks act as a paradoxical pedagogical catalyst that can either hinder or help the student. For students with low motivation, the tasks are experienced as technical friction. In contrast, for intrinsically motivated students, they serve as necessary, metacognitive "speed bumps". They force the student out of a pure "gaming" mode, requiring textbook lookups and the application of Excel to validate actions performed in the simulator. A common comment from those who report that they find the narrative tasks to be disrupting is that they would have preferred it if the data in the assignments were more directly tied to their individual game state. This finding suggests that external, theoretical assignments within the simulator should be retained, but their relevance must be clearly communicated to prevent a loss of motivation among less engaged students.

Considering the Conversational Framework, this illustrates how the simulator activity provides opportunities for switching between different learning types. The narrative tasks prompt switching between practice and acquisition (and, for some students, also discussion and collaboration). However, this is contingent on the motivational profile of the student. We think this serves as a good example of metamotivational fit (Miele and Scholer, 2018). The incorporation of the simulator in the course design provides sufficiently motivated students with opportunities to perform cycles of different learning types. This links the simulator to other course materials in a meaningful way that is appropriate for their motivation. On the other hand, for students who lack the motivation to engage in this form of task-switching, it contributes to an experience of less meaningful activity.

In our current dataset, we see one difference from the previous studies: Students with the highest level of reflection reported slightly lower perceived learning than those with a medium level of reflection. This is what we expected in our original study. The qualitative data offer a rational explanation: Students with the deepest business acumen seem more capable of identifying the simulator's limitations (e.g., lack of HR factors, nuanced marketing, or inflation). Their deeper reflection results in a more self-critical and sober assessment of their own perceived learning. Consequently, this can be seen as not a sign of diminished learning but rather as pointing towards analytical maturity. In the text responses about the narrative assignments, many highlight them as more important for learning than the simulator itself. It is possible that the students who master switching between the different tasks and modes of learning also become more aware of the limits of the simulator.

5. Conclusion

This study demonstrates that in our asynchronous, online educational context, the students' internal psychological states act as the primary engine for learning, while external pedagogical designs serve as indirect catalysts. Addressing our primary research question, we conclude that intrinsic motivation, engagement and reflection are the prominent direct drivers of self-reported perceived learning from the use of a business simulator in our context. These internal states override demographic factors, establishing the simulator as a gender-neutral learning arena where effort, reflection and motivation facilitate learning.

The main contribution of this article is that it corrects and validates our Tiers of Engagement framework. Our statistical analysis has resulted in the same overall findings as in the previous study for how intrinsic motivation, engagement and depth of reflection drive reported perceived learning. We did not find the same statistical significance for the inclusion of narrative assignments and collaboration. Through statistical analyses, we have examined how collaboration serves as a catalyst for motivation, engagement and reflection. The narrative tasks are shown to serve as a catalyst for reflection on learning. Unlike other studies, we do not find any significant impact on perceived learning from gender in our context.

Our mixed-methods analysis allows us to peek behind the metrics and examine the qualitative data from the perspective of the Conversational Framework. This sheds light on how students experience the simulator as a part of the overall pedagogical design. For instance, opportunities for motivating iterations of different activity types are important. By facilitating successful task-switching between practical gameplay and theoretical acquisition, students can break out of a superficial "gaming" mode, enabling a strong metamotivational fit. While the quantitative data showed that the narrative assignments and the optional collaboration did not directly increase perceived learning, the qualitative evidence highlights their role as a catalyst. For the highly motivated students, this offers increased potential for learning, while the less motivated students experience disruption. Similarly, while students are limited in their ability to participate in discussions and collaboration, it can function as a powerful social buffer; for those who engage in it, it resolves technical and academic friction, thereby sustaining the engagement, motivation and reflection necessary to promote learning.

In continuing development of the overall pedagogical design, we see it as important to address these findings. There might be further opportunities in adding more variety to the types of learning activities offered in the course, but most importantly to explore more ways of reducing the frustration or disruption experienced when switching between "playing the simulator" and using it as a learning tool. Hopefully this can reduce the chance of the narrative assignments being experienced as "metacognitive speedbumps" that brings the learning process to a halt. With some adjustments, they might instead become "metacognitive ramps" that assists students in jumping to a higher tier of engagement.

Unlike in our previous studies, we identified a non-linear relationship where students exhibiting the highest levels of reflection reported slightly lower perceived learning than their moderately reflecting peers. This was our initial expectation when we started working on our framework. From our analysis of the qualitative data, we do not see this as a deficit in learning, but potential analytical maturity. The most deeply engaged students develop the business acumen necessary to identify the simulator's inherent limitations, resulting in a more sober, self-critical assessment of their own learning. As mentioned, we see that the higher reflecting students call attention to the importance of the narrative assignments, some highlighting that they found them more important than the simulator. Continuing the "metacognitive ramp" analogy, we see potential in expanding the ramp to serve as a bridge between practice and theoretical learning. By refining the combination of learning activities, we can strengthen the learning opportunities that arise from contrasting a simplified simulation with real-world contexts. In light of this, instructors and researchers need to be aware that when measuring perceived learning, there can be variables of internal psychological states that disrupt the student's own perception of learning, as we see with depth of reflection.

Acknowledgements

We are grateful for feedback on the article from our colleague Eirik Heen, and the anonymous reviewers.

Ethics Declaration: The data was collected with informed consent and anonymously.

AI Declaration: We made use of a LLM (Google Gemini Pro) in the analysis of the data, as described in the analysis section. Furthermore an LLM (Google Gemini Pro) was also used in the writing process to structure the text, improve grammar, and improve the readability of the article. All text has been manually quality assured by the authors to confirm that all text is representative of the authors intended message.

References

- Almeida, F., 2018. Strategies to perform a mixed methods study. *European Journal of Education Studies*, 5(1), pp. 137–151.
- Butler, F. C., Mullen, D. and Wheatley, K. K., 2025. Business simulation performance: Who performs better? *Journal of Education for Business*, 100(1), pp. 35–44.
- Costello, A. B. and Osborne, J. W., 2005. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, 10(1), article 7.
- Deci, E. L. and Ryan, R. M., 2000. The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), pp. 227–268.
- Delacre, M., Leys, C., Mora, Y. L. and Lakens, D., 2019. Taking parametric assumptions seriously: Arguments for the use of Welch's F-test instead of the classical F-test in one-way ANOVA. *International Review of Social Psychology*, 32(1).
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. and Strahan, E. J., 1999. Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), pp. 272–299.
- Field, A., 2018. *Discovering statistics using IBM SPSS statistics*. 5th ed. London: Sage.
- Games, P. A. and Howell, J. F., 1976. Pairwise multiple comparison procedures with unequal N's and/or variances: A Monte Carlo study. *Journal of Educational Statistics*, 1(2), pp. 113–125.
- Gonyea, R. M., 2005. Self-reported data in institutional research: Review and recommendations. *New Directions for Institutional Research*, 2005(127), pp. 73–89.
- Guttormsen, M. A. B. and Weines, J., 2024. Tiers of Engagement: Achieved Learning From Business Simulations Reflected in Economics Students' Experiences. *European conference on games based learning*, 18 (1), pp. 320–328.
- Guttormsen, M. and Weines, J., 2025. Tiers of Engagement II: Profiling the Characteristics of User Engagement and Motivational Tiers in Business Simulators. *European conference on games based learning*, 19 (1), pp. 357–367.
- Laurillard, D., 2012. *Teaching as a design science: Building pedagogical patterns for learning and technology*. London: Taylor & Francis Group.
- Miele, D. B. and Scholer, A. A., 2018. The role of metamotivational monitoring in motivation regulation. *Educational Psychologist*, 53(1), pp. 1–21.
- R Core Team, 2025. *R: A language and environment for statistical computing* (Version 4.5) [Computer software]. Vienna: R Foundation for Statistical Computing.
- Ryan, R. M. and Deci, E. L., 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), pp. 54–67.
- Schwartz, R. N. and Plass, J. L., 2019. Types of engagement in learning with games. In: J. L. Plass, R. E. Mayer and B. D. Homer, eds. *Handbook of game-based learning*. Cambridge, MA: The MIT Press, pp. 53–80.
- The jamovi project, 2025. *jamovi* (Version 2.7) [Computer software].

Appendix 1: Coding Criteria

Table 3: Criteria for coding reflection on learning

Level of reflection on learning indicators:	How the students describe their experience with the simulator, their perceived learning from it, and insight regarding their learning process. We differentiate between how detailed and deep their thoughts about the simulator and their own learning are as described below.
High (3):	Reflects on what they have learned, but also how they learned, and what could facilitate more learning. Relating their learning to larger contexts.
Medium (2):	Reflects on what they have learned.
Low: (1)	Reflects on what tasks they have done and if they have learned/not learned (e.g. answering "yes" or "no" without further specifications).

Table 4: Criteria for coding level of intrinsic motivation for the simulator

Level of intrinsic motivation indicators:	How the students describe their experience with the simulator with regards to their feelings of joy and interest towards working with the simulator.
High (H):	Answered "I thought the simulator was very fun and interesting".
Medium (M):	Answered "I thought the simulator was somewhat fun and interesting".
Low: (L)	Answered "I thought the simulator was a neither fun nor interesting", "I thought the simulator was a little boring/not interesting" or "I thought the simulator was very boring/uninteresting"

Table 5: Criteria for coding learning from narrative assignments

Level of learning from narrative assignments indicators:	How the students describe their perceived learning or usefulness of the narrative assignments put in by the course instructor.
High (3):	Comments indicating great learning value, such as: "Very helpful or very educational", etc.
Medium (2):	Comments indicating some learning value, such as: "OK", "a bit helpful", etc.
Low: (1):	Comments indicating no or negative learning value, such as: "disruptive", "confusing", "bad" etc.

Table 6: Criteria for coding degree of collaboration

Level of collaboration or discussion	How respondents described their collaboration or discussion with others in regards to the simulator
High (3)	High level of discussion or collaboration with others
Some (2)	Some collaboration or discussion with others
None (1)	Did not collaborate or discuss / Not applicable