

First-Month LMS Behaviours Associated with Success in an AI-Permissive Software Engineering Course

1st Andrea Vázquez-Ingelmo
Department of Computer Science
University of Salamanca
Salamanca, Spain
andreavazquez@usal.es

2nd M. Puerto Paule-Ruiz
Department of Computer Science
University of Oviedo
Oviedo, Spain
paule@uniovi.es

3rd Francisco José García-Peñalvo
Department of Computer Science
University of Salamanca
Salamanca, Spain
fgarcia@usal.es

4th Alicia García-Holgado
Department of Computer Science
University of Salamanca
Salamanca, Spain
aliciagh@usal.es

Abstract—This extended abstract profiles the highest-performing behavioural cluster during the first milestone of a Software Engineering course governed by an explicit artificial-intelligence (AI) usage policy. Moodle traces from two post-AI cohorts ($n = 146$) yielded three clusters: an *integrator* group ($n = 32$), an *intermediate* group ($n = 73$), and an *at-risk* group ($n = 41$). Integrators achieved a higher mean milestone grade than at-risk students (6.37 versus 4.81; Hedges' $g = 0.92$) and a higher pass rate (71.9% versus 44.0%). Their profile combined access volume and anticipation: they accessed theoretical materials, task specifications, rubrics, and modelling tools more often, and reached theoretical materials about ten days earlier. A logistic model based on theoretical-material volume, task volume, and task first-access timing separated integrators from the remaining students with an ROC AUC of 0.86. The results support first-month onboarding that introduces rubrics and modelling tools early, triggers early task access, uses analytics for positive recognition, and treats forum activity as a complement rather than a substitute for task- and rubric-anchored engagement.

Index Terms—artificial intelligence, computing education, learning management systems, student profiles, self-regulated learning, AI policy, onboarding

I. INTRODUCTION

Generative AI has encouraged institutions to adopt explicit policies that allow its use when students disclose and document it [1], [2]. In computing courses, such policies have been associated with changes in engagement and grades [3], [4]. In the course studied here, students form distinct behavioural clusters during the first milestone (M1), with the highest-performing group averaging 6.37 on a 0–10 scale and the lower-performing group averaging 4.81 [3].

This study asks what the top-performing students do differently during the first month and whether that profile can be approximated using a small set of interpretable LMS features.

The purpose is to identify concrete practices that can inform early onboarding.

II. CONTEXT AND METHOD

Software Engineering I is a second-year core course at the University of Salamanca organised through project-based learning. Since 2023–24, students have been permitted to use AI tools provided that they document prompts, reasoning, and outputs in their portfolios. The analysis includes the complete 2023–24 and 2024–25 cohorts ($n = 65$ and $n = 81$), giving $n = 146$ students for behavioural analyses and $n = 129$ for grade comparisons; 17 students had no recorded M1 grade.

Moodle activity was summarised by student and resource type using access volume and timing indicators. Missing behavioural values were set to zero when a resource was never accessed; grades were not imputed. Following the previous analysis of these cohorts [3], features were standardised, projected onto two principal components, and clustered using K-means with $K = 3$. The highest-grade cluster was labelled *integrator* (C0), with the others labelled *at-risk* (C1) and *intermediate* (C2). Grade differences were tested with ANOVA, Tukey HSD, and Hedges' g ; behavioural features were compared through univariate ANOVAs and medians. Logistic regression estimated integrator membership from theoretical-material volume, task volume, and task first-access timing.

III. RESULTS

The clusters contained 32 integrators, 73 intermediate students, and 41 at-risk students. Grade differences were significant overall ($F(2, 126) = 6.42$, $p = 0.002$, $\eta_p^2 = 0.093$). The main contrast was between integrators and at-risk students: a 1.55-point difference ($p = 0.001$; Hedges' $g = 0.92$). The at-risk and intermediate clusters also differed significantly

TABLE I
SELECTED OUTCOMES AND ACTIONABLE LMS MEDIAN

Measure	Integrator	At-risk
Students, n	32	41
M1 grade, mean (SD)	6.37 (1.67)	4.81 (1.66)
Pass rate	71.9%	44.0%
Theoretical-material TNA	91	36
Task TNA	21	1
Rubric TNA	4.5	0
Tool TNA	7.5	0
Theoretical ATFA, days	12.2	22.1

($g = -0.58$, $p = 0.036$), whereas the integrator–intermediate contrast was not significant.

The integrator profile combined volume and anticipation (Table I). Integrators accessed theoretical materials 2.5 times more often than at-risk students and task specifications 21 times more often. The median integrator opened the rubric and modelling tools repeatedly, whereas the median at-risk student never opened either during M1. Integrators also reached theoretical materials about ten days earlier. The most discriminative variables included tool TNA, theoretical-material ATFA, rubric TNA, task TNA, theoretical-material TNA, and rubric ATFA.

Forum activity did not distinguish the integrator cluster. Announcements are forwarded automatically to institutional email, so forum TNA represents deliberate returns to Moodle rather than exposure to essential information. Deliberate forum browsing without corresponding rubric, task, or tool engagement was not associated with a grade advantage.

The logistic model achieved an ROC AUC of 0.857. A conservative rule—theoretical-material TNA ≥ 91 and task ATFA ≤ 10.1 days—had sensitivity 0.28, specificity 0.97, and positive predictive value 0.75. It identified only a subset of integrators, but most flagged students were true integrators.

IV. DISCUSSION AND IMPLICATIONS

The integrator profile is more specific than generic high engagement: it reflects early, repeated interaction with the artefacts that define the deliverable. In self-regulated-learning terms, this is consistent with a stronger forethought phase [5], [6]. Repeated rubric access may also have an AI-specific role: students could use assessment criteria to formulate prompts and evaluate AI-generated outputs. Because individual AI use was not measured, this remains a hypothesis and may instead reflect general conscientiousness or assessment strategy.

The profile supports four first-month actions from the full study: introduce the rubric, AI-documentation requirements, and modelling tools in week 1; require an early task that opens and annotates the specification; use the high-specificity rule for positive recognition; and pair forum scaffolds with rubric- and task-anchored activities.

The study is observational and limited to one course at one institution. LMS traces omit synchronous sessions, peer collaboration, offline study, and direct AI-tool use. Replication

across disciplines and institutions, together with direct measurement of AI use, is needed.

V. CONCLUSION

Across two post-AI cohorts, the highest-performing cluster was characterised by early and repeated engagement with theoretical materials and the artefacts defining the deliverable. This profile was associated with a 1.55-point advantage over the at-risk cluster, a nearly 28-percentage-point pass-rate gap, and an interpretable three-feature model with an ROC AUC of 0.86. The findings translate the behavioural profile into concrete onboarding guidance while remaining cautious about causal interpretation.

ACKNOWLEDGMENT

This work was supported by the Ministry of Science, Innovation and Universities and co-funded by the European Regional Development Fund (ERDF) under project PID2024-155586OB-I00 (MCIU/AEI/10.13039/501100011033). Additional support was provided by the Government of the Principality of Asturias through grant GRU-GIC-24-070.

REFERENCES

- [1] F. Miao and W. Holmes, *Guidance for Generative AI in Education and Research*. Paris, France: UNESCO, 2023. [Online]. Available: <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- [2] M. Alier, F. J. García-Peñalvo, M. J. Casañ, J. A. Pereira, and F. Llorens-Largo, *Safe AI in Education Manifesto*, version 0.4.0, 2024. doi: 10.5281/zenodo.13821788.
- [3] A. Vázquez-Ingelmo, M. P. Paule-Ruiz, F. J. García-Peñalvo, and A. García-Holgado, “Learning to learn with AI: how AI policy reshapes learning trajectories in software engineering education,” *Behaviour & Information Technology*, 2026, doi: 10.1080/0144929X.2026.2676747.
- [4] J. Prather, B. Reeves, J. Leinonen, S. MacNeil, A. S. Randrianasolo, B. A. Becker, B. Kimmel, J. Wright, and B. Briggs, “The widening gap: The benefits and harms of generative AI for novice programmers,” in *Proc. 2024 ACM Conference on International Computing Education Research – Volume 1 (ICER 2024)*, 2024. doi: 10.1145/3632620.3671116.
- [5] I. Molenaar, “The concept of hybrid human–AI regulation: Exemplifying how to support young learners’ self-regulated learning,” *Computers and Education: Artificial Intelligence*, vol. 3, Art. no. 100070, 2022. doi: 10.1016/j.caeai.2022.100070.
- [6] B. J. Zimmerman, “Attaining self-regulation: A social cognitive perspective,” in *Handbook of Self-Regulation*, M. Boekaerts, P. R. Pintrich, and M. Zeidner, Eds. San Diego, CA, USA: Academic Press, 2000, pp. 13–39. doi: 10.1016/B978-012109890-2/50031-7.