Supplementary material

Post-pandemic challenges in higher education: learning preferences, performance and dropout in a first-semester chemistry course

Nicolás Pérez, Lorena Martínez, Natalia Alvarez, Lucía Otero, Nicolás Veiga*, Julia Torres*

Área Química Inorgánica, DEC, Facultad de Química, Universidad de la República, General Flores 2124, Montevideo, Uruguay

*corresponding authors: nveiga@fq.edu.uy, jtorres@fq.edu.uy

Supplementary tables

Table S1. Summary of assessment practices before, during, and after the COVID-19 pandemic.

Period	Test Format	Modality	Supervision	Test Composition and Design	Evaluation Criteria	Notes
Pre-pandemic (2017-2019)	Two multiple-choice tests	In-person	On site by instructors	Multiple-choice questions (5 options, 1 correct). Questions drawn from a validated database covering theoretical and practical aspects with predefined proportions of difficulty and time demand	Final grade = % of correct answers – 20% (guessing correction). Course passed with ≥50% of total possible score	Standard assessment format: all students of a cohort take identical tests; no additional graded activities
Early pandemic (2020)	Same as pre-pandemic	One online and one in-person	On site or online in small Zoom groups by instructors	Same as pre-pandemic	Same as pre-pandemic	First evaluation was conducted remotely
Late pandemic (2021)	Same as pre-pandemic	Online	Online in small Zoom groups by instructors	Same as pre-pandemic	Same as pre-pandemic	Full evaluation was conducted remotely
Post-pandemic (2022-2023)	Same as pre-pandemic	Same as pre-pandemic	Same as pre-pandemic	Same as pre-pandemic	Same as pre-pandemic	Return to standard assessment format

Table S2. Variables selected for this study and their descriptive parameters.[a]

Group of variables		Variable	Description	n	mean	sd	min	max	Detail	Period
General descriptors of students		GDR	Gender	4511	0.32	0.47	0	1	(Female=0; Male=1)	2017-2023
		Y	Year of enrolment	4511	---	---	2017	2023	---	---
		RP	Repeating the course	4511	0.23	0.42	0	1	(NO=0; YES=1)	2017-2023
		ON	Taking the flexible modality	4511	0.08	0.27	0	1	(NO=0; YES=1)	2017-2023
		G1	Previous academic performance	4277	34.55	23.67	0	100	First evaluation in marks %	2017-2023
		DR	Dropout	4511	0.19	0.39	0	1	Students taking only first evaluation (NO=0; YES=1)	2017-2023
		GCI	Final academic performance	4511	29.06	22.98	0	100	In-presence multiple-choice tests marks (%)	2017-2023
Access / use / usefulness	Theory	T_U	Zoom synchronous lectures use frequency	389	2.28	1.10	1	4	Survey (Likert 1-4)	2022
		T_Uf	Zoom synchronous lectures perceived usefulness	383	3.29	1.16	1	5	Survey (Likert 1-5)	2022
		TV_A	Access to recorded lectures	2816	44.73	33.12	0	100	YouTube access to videos	2020 - 2023
		TV_U	Lectures videos use frequency	391	3.11	1.01	1	4	Survey (Likert 5111-4)	2022
		TV_Uf	Lectures videos perceived usefulness	390	4.15	1.04	1	5	Survey (Likert1-5)	2022
		AV_A	Access to animated YouTube videos	2816	17.13	25.20	0	100	YouTube access to videos	2020 - 2023
		AV_U	Animated YouTube videos use frequency	390	1.92	0.95	1	4	Survey (Likert1-4)	2022
		AV_Uf	Animated YouTube videos perceived usefulness	387	3.14	1.08	1	5	Survey (Likert1-5)	2022
		RM_A	Access to recommended reading material	2816	40.57	34.34	0	100	VLE registered access	2020 - 2023
		RM_U	Recommended reading material use frequency	391	3.04	0.96	1	4	Survey (Likert1-4)	2022
		RM_Uf	Recommended reading material perceived usefulness	386	4.14	0.98	1	5	Survey (Likert1-5)	2022
	Practice	PV_A	Access to recorded practical classes	2816	29.99	30.67	0	100	YouTube access to videos	2020 - 2023
		PV_U	Recorded practical classes use frequency	391	2.42	1.08	1	4	Survey (Likert1-4)	2022
		PV_Uf	Recorded practical classes perceived usefulness	386	3.85	1.05	1	5	Survey (Likert1-5)	2022

		CL_Uf	Solving clues perceived usefulness	1041	4.02	0.89	1	5	Survey (Likert1-5)	2017 2020 2022
		FK_Uf	Feedback answers perceived usefulness	1044	4.17	0.85	1	5	Survey (Likert1-5)	2017 2020 2022
		KW_Uf	Daily-life curiosities “know-what" sections perceived usefulness	1056	3.43	0.95	1	5	Survey (Likert1-5)	2017 2020 2022
		KM_Uf	Higher-level “know-more" sections perceived usefulness	1029	3.42	0.95	1	5	Survey (Likert1-5)	2017 2020 2022
		Hy_Uf	History sections perceived usefulness	1026	3.19	0.94	1	5	Survey (Likert1-5)	2017 2020 2022
	Student-educator interaction	TS_U	Time for student queries use frequency	391	1.95	0.96	1	4	Survey (Likert1-4)	2022
		TS_Uf	Time for student queries perceived usefulness	386	3.38	0.98	1	5	Survey (Likert1-5)	2022
		EF_A	Access to electronic forum	2816	0.83	0.38	0	1	(NO=0; YES=1)	2020 - 2023
		EF_U	Electronic forum use frequency	389	1.92	0.86	1	4	Survey (Likert1-4)	2022
		EF_Uf	Electronic forum perceived usefulness	386	3.32	0.95	1	5	Survey (Likert1-5)	2022
	Self-assessment	M_U	Mocks use frequency	389	0.92	0.27	0	1	Survey (NO=0; YES=1)	2022
		M_Uf	Mocks perceived usefulness	389	4.52	0.81	1	5	Survey (Likert1-5)	2022

[a] n = sample size, sd = standard deviation, min = minimum value, max = maximum value.

Table S3. Temporal evolution of students' descriptors expressed as student percentages.[a]

year	n	GDR = 1 (%)	DR = 1 (%)	RP = 1 (%)	ON = 1 (%)	GCI (average in marks %)	Percent of students above passing grade	Period
2017	563	33.0	13.9	20.8	6.4	33.6	21.5	pre-pandemic
2018	560	33.9	14.8	20.0	4.6	33.3	23.0	
2019	572	33.2	14.2	18.5	5.6	28.9	20.8	
2020	718	31.9	34.1	25.9	10.5	22.8	16.2	pandemic
2021	739	31.1	20.2	26.1	6.8	30.2	21.5	
2022	637	30.0	13.5	27.3	9.4	28.6	18.2	post-pandemic
2023	722	32.6	18.6	23.5	10.9	27.9	17.3	
Total sample	**4511**	**32.2**	**19.0**	**23.5**	**7.9**	**29.1**	**19.6**	

[a] GDR = 1 (male student), RP = 1 (student repeating the course), ON = 1 (students that chose online modality of the course), DR = 1 (students dropping out after the first test).

Table S4. Average percentage variation ($\bar{x}$) and confidence intervals (*CI*) for general descriptors of the students across different selected periods.

Comparison	GDR = 1 (%)[a]		RP = 1 (%)[a]		ON = 1 (%)[a]		DR = 1 (%)[a]		GCI (average; %)[b]	
	$\bar{x}$	*CI* 95%	$\bar{x}$	*CI* 95%	$\bar{x}$	*CI* 95%	$\bar{x}$	*CI* 95%	$\bar{x}$	*CI* 95%
2020 to pre-pandemic	-1.5	[-6.7, 3.7]	**6.1***	**[1.6, 10.7]**	**4.9***	**[2.1, 7.7]**	**19.8***	**[15.4, 24.3]**	**-9.1***	**[-11.8, -6.4]**
2021 to 2020	-0.8	[-6.9, 5.3]	0.2	[-5.5, 6.0]	**-3.7***	**[-7.3, 0.0]**	**-14.0***	**[-19.8, -8.1]**	**7.5***	**[4.3, 10.6]**
Post-pandemic to 2021	0.2	[-5.1, 5.5]	-0.8	[-5.8, 4.2]	**3.5***	**[0.2, 6.7]**	-4.0	[-8.3, 0.4]	-2.0	[-4.6, 0.6]
Post-pandemic to pre-pandemic	-2.0	[-6.3, 2.2]	**5.5***	**[1.8, 9.3]**	**4.7***	**[2.3, 7.1]**	1.9	[-1.3, 5.2]	**-3.6***	**[-5.8, -1.6]**

* Statistical significance at 95% level of confidence. [a] Two-sample Proportion z-test with Bonferroni correction per variable. [b] Games-Howell test.

Table S5. Average variation in the percentage of resources accessed ($\underline{x}$) and their confidence intervals (*CI*) across different selected periods.

Comparison	TV_A (%)[a]		AV_A (%)[a]		PV_A (%)[a]		RM_A (%)[a]		EF_A = 1 (%)[b]	
	$\underline{x}$	*CI* 95%	$\underline{x}$	*CI* 95%	$\underline{x}$	*CI* 95%	$\underline{x}$	*CI* 95%	$\underline{x}$	*CI* 95%
2021 – 2020	**23.8***	**[19.6, 28.0]**	**22.2***	**[19.0, 25.4]**	**20.7***	**[16.6, 24.9]**	**22.1***	**[17.5, 26.8]**	**8.6***	**[5.1, 12.1]**
2022 – 2021	-4.1	[-8.7, 0.4]	**-9.0***	**[-12.7, -5.2]**	**-18.1***	**[-22.5, -13.7]**	**-5.7***	**[-10.7, -0.7]**	**-4.8***	**[-8.0, -1.6]**
2023 – 2022	-3.9	[-8.4, 0.6]	**-8.8***	**[-12.1, -5.4]**	-1.9	[-5.8, 2.0]	**-7.9***	**[-12.4, -3.4]**	**-34.5***	**[-40.5, -28.4]**

* Statistical significance at 95% level. [a] Games-Howell test. [b] Two-sample Proportion z-test with Bonferroni correction per variable. TV_A (videos of theoretical classes), AV_A (animated complementary videos), PV_A (recorded practical worked examples), RM_A (recommended reading material), EF_A (electronic forum).

Table S6. Spearman's rank correlation coefficient between the access (_A) and declared use frequency (_U) of the available resources in 2022.

Resource	Correlation coefficient (_A *vs*. _U)
TV	0.49*
AV	0.25*
PV	0.45*
RM	0.40*
EF	0.10

* Statistical significance at 95% level.

Table S7. Spearman's rank correlation coefficient between the declared use frequency (_U) and perceived usefulness (_Uf) of the available resources in 2022.

Resource	Correlation coefficient (_U *vs*. _Uf)
TV	0.60*
AV	0.73*
PV	0.74*
RM	0.71*
TS	0.56*
EF	0.60*
M	0.47*

* Statistical significance at 95% level.

Table S8. Temporal variation of the declared use and usefulness of the interactive materials' resources (τ).[a]

Comparison	CL_Uf	FK_Uf	KW_Uf	KM_Uf	Hy_Uf
2020 – 2017	**0.12***	**0.17***	-0.04	**0.17***	0.05
2022 – 2020	-0.07	**-0.18***	**-0.26***	**-0.35***	**-0.29***
2022 – 2017	0.04	-0.03	**-0.29***	**-0.21***	**-0.24***

[a] τ = Kendall's tau-b. * Statistical significance at 95% level of confidence (Bonferroni correction per variable).

Table S9. Multiple linear regression analysis (*y* = GCI, average in marks %) on samples containing students grouped by previous academic performance (G1 thirds) or repeater students.

	Highest G1 third	Medium G1 third	Lowest G1 third	Repeater students
Number of students	36	151	195	95
Explained variance	83%	59%	59%	89%
Variables	*Regression coefficients*[a]			
Gender (GDR)	---	---	-1.66	-1.60
Online coursing (ON)	**7.80****	---	-3.08	3.26
Course repetition (RP)	4.71	---	---	---
First test mark (G1)	**1.01*****	**1.02*****	**0.84*****	**0.87*****
Course drop-out (DR)	---	**-17.3***	**-8.65*****	**-12.7*****
TV_U	---	---	0.75	---
AV_U	-1.30	---	---	1.34
PV_U	---	-0.66	---	**-2.03***
RM_U	---	---	---	**1.83***
TS_U	---	**1.53***	---	**2.20****
EF_U	-1.01	**-2.16****	-0.72	**-1.95***
M_U	**12.5***	---	-3.29†	-2.60
FK_Uf	**2.60****	---	---	---
KW_Uf	---	---	---	**3.78****
KM_Uf	---	---	---	2.11
Hy_Uf	---	---	---	**-5.62*****

[a] Statistical significance: $p<0.001$ (***), $p<0.01$ (**), $p<0.05$ (*), $p<0.1$ (†).

Supplementary figures

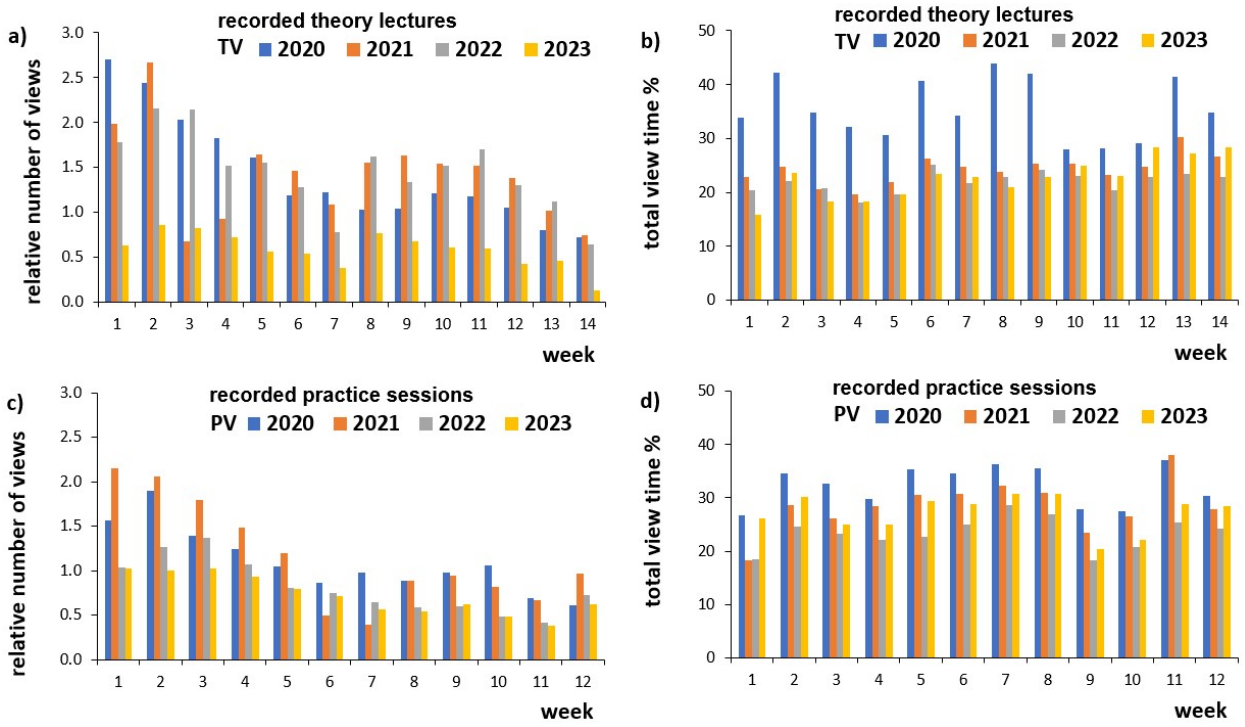

Figure S1. Temporal evolution of the relative number of views—proportion of the total number of views relative to enrolled students—(a, c) and the total view time %— average percentage of the video's total duration watched by viewers —(b, d) for the videos of theory lectures (TV) (a, b) and practice sessions (PV) (c, d).

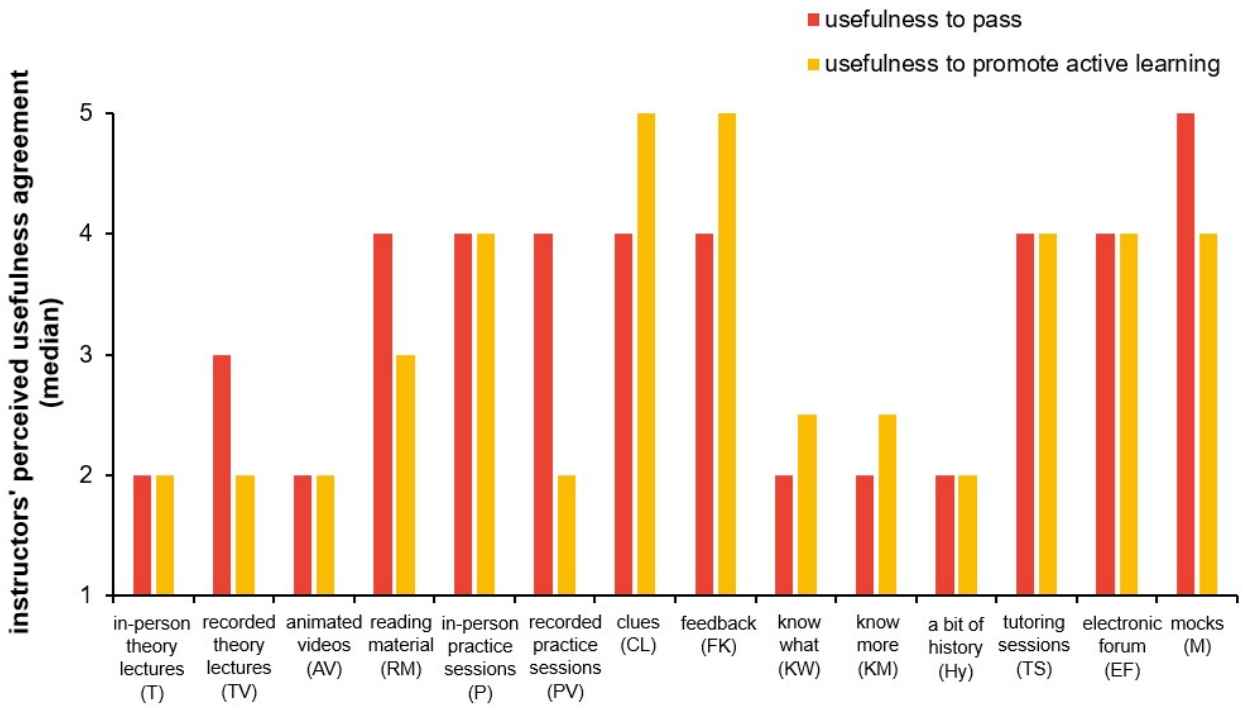

Figure S2. Instructors' survey results: median values obtained for the degree of agreement with regard to the usefulness to pass tests or to promote active learning (5: I totally agree, 4: I agree, 3: I neither agree nor disagree, 2: I disagree, and 1: I totally disagree).

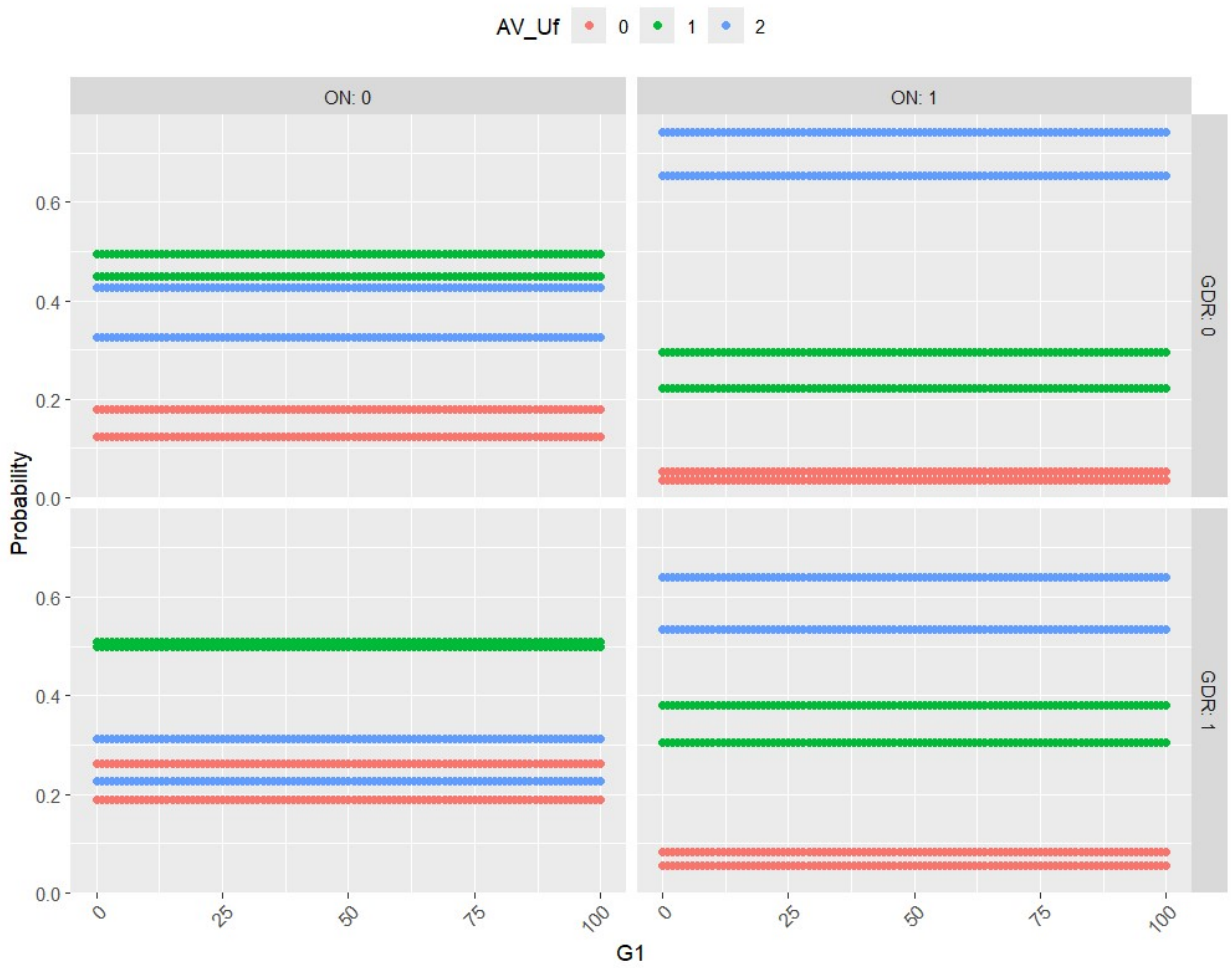

Figure S3. Probability of perceiving the complementary animated videos, AV, useful as a function of previous academic performance (G1), gender (GDR) and online coursing (ON).

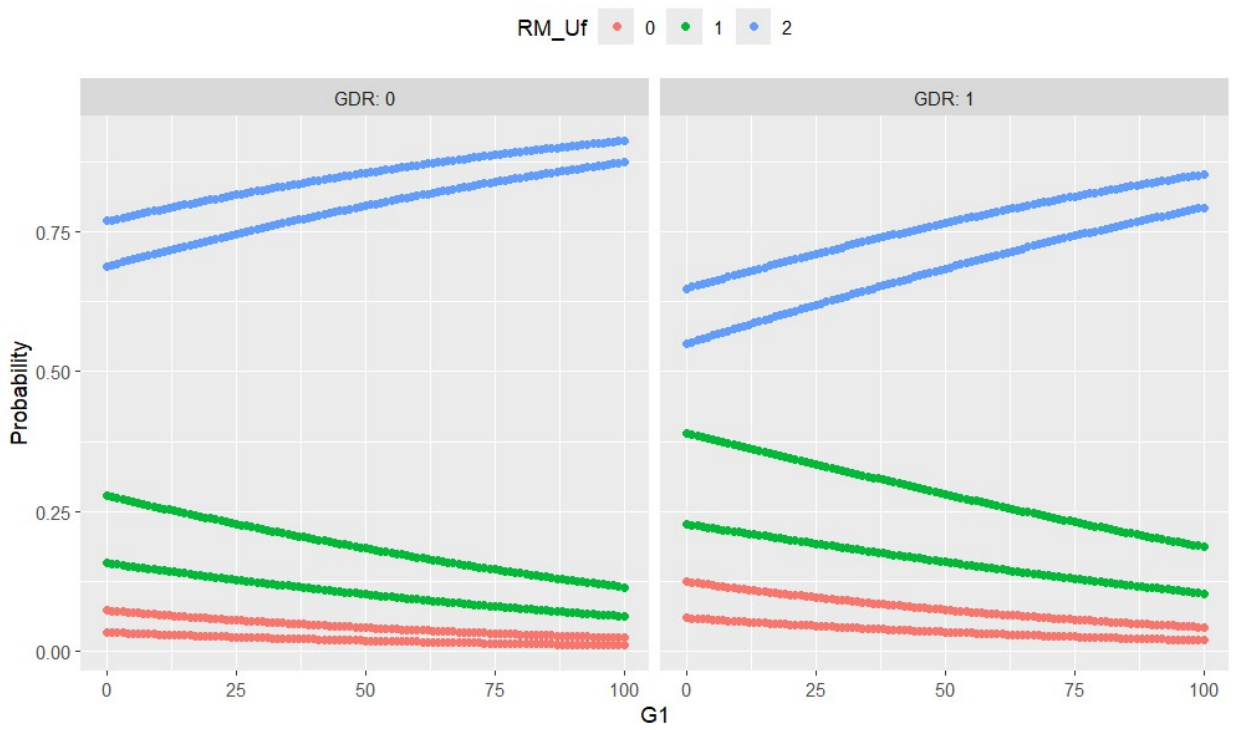

Figure S4. Probability of perceiving the recommended reading material, RM, useful as a function of previous academic performance (G1) and gender (GDR).

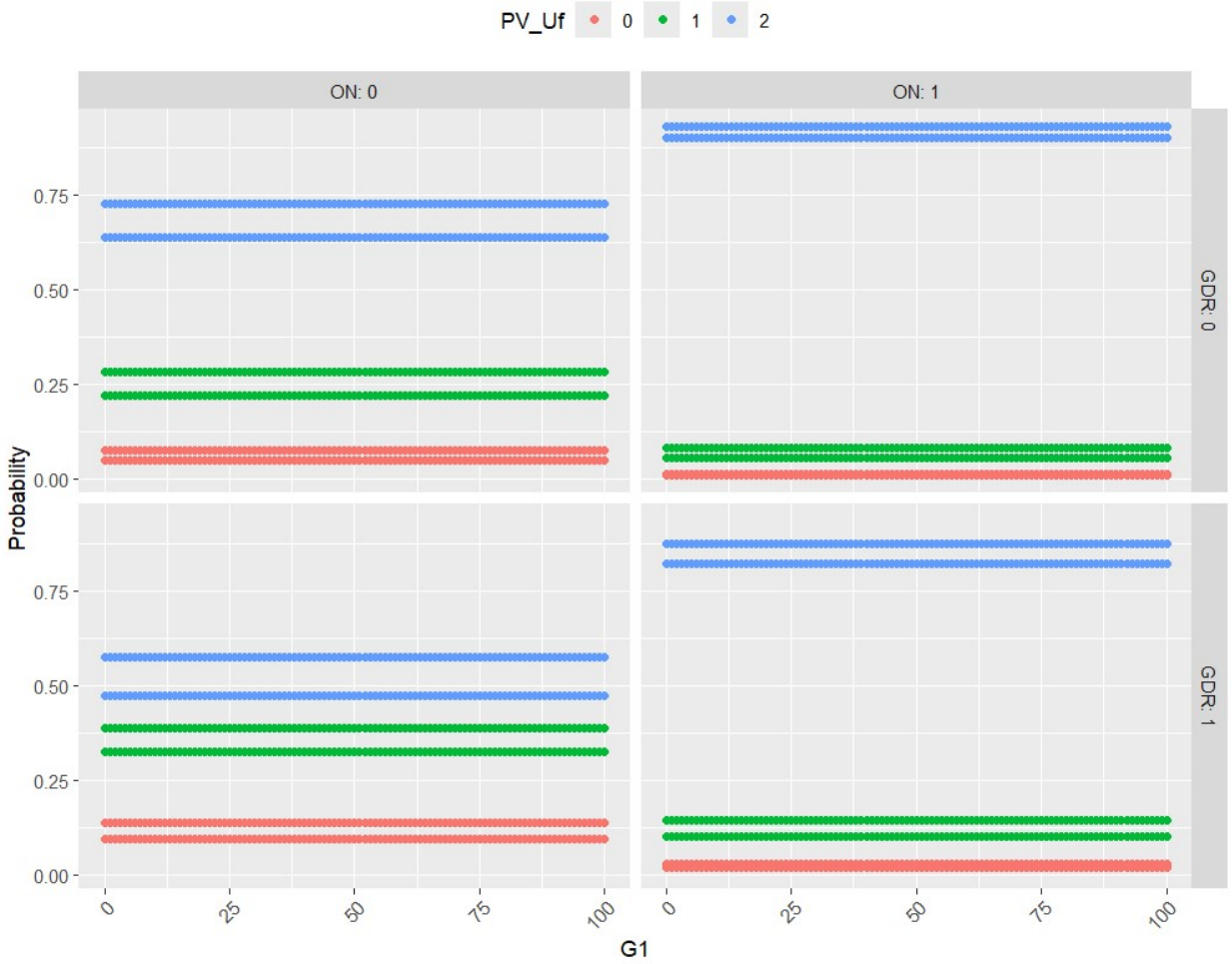

Figure S5. Probability of perceiving the recorded practical classes, PV, useful as a function of previous academic performance (G1), gender (GDR) and online coursing (ON).

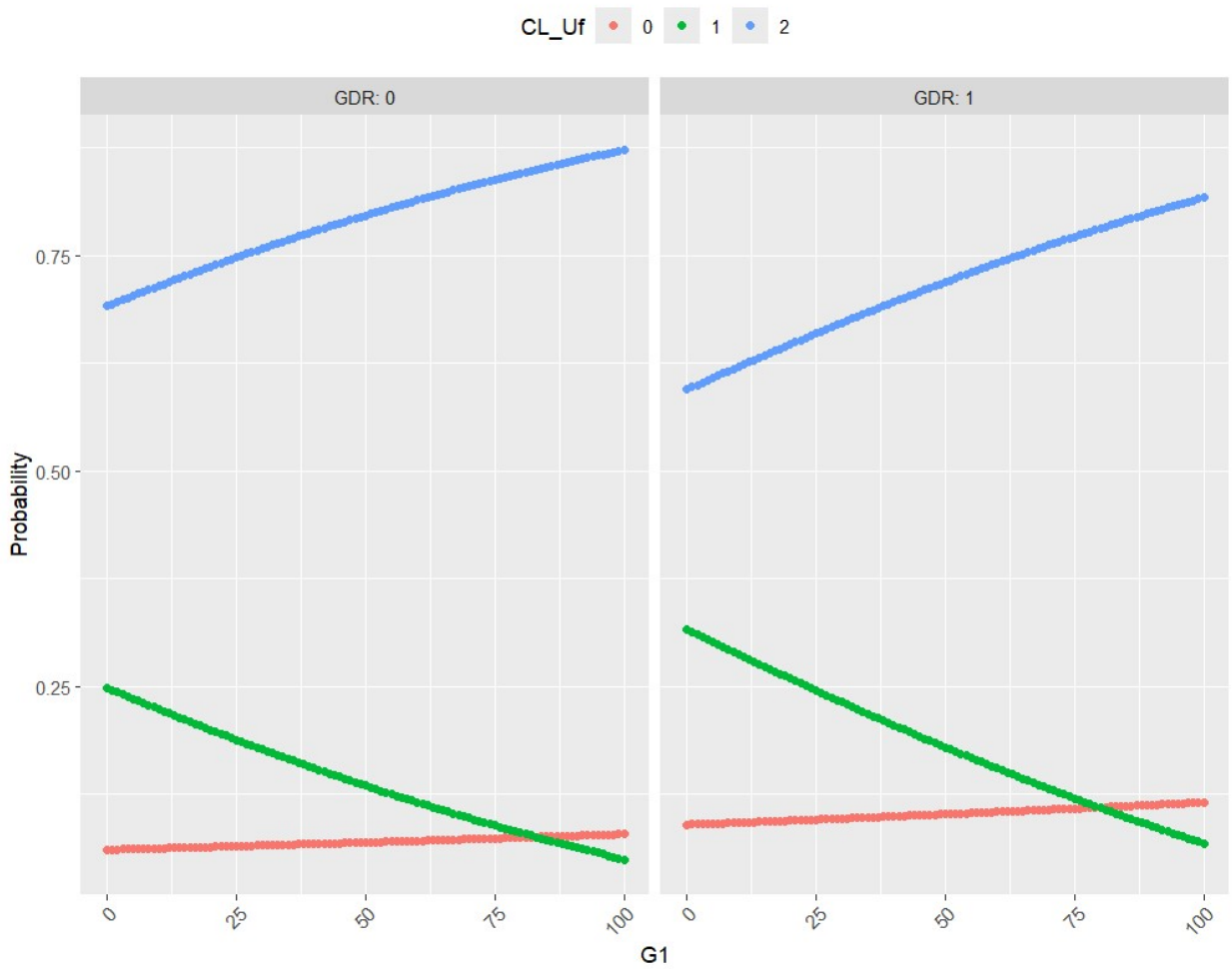

Figure S6. Probability of perceiving the solving clues, CL, useful as a function of previous academic performance (G1) and gender (GDR).

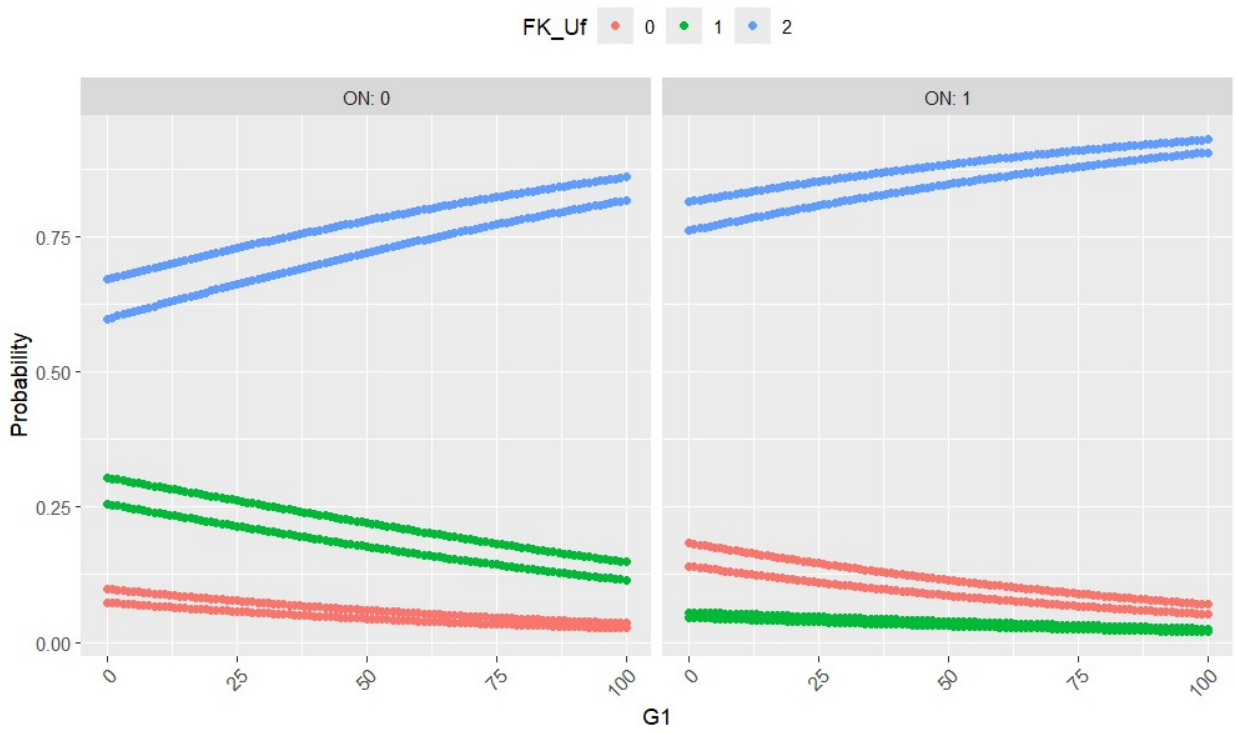

Figure S7. Probability of perceiving the feedback answers, FB, useful as a function of previous academic performance (G1) and online coursing (ON).

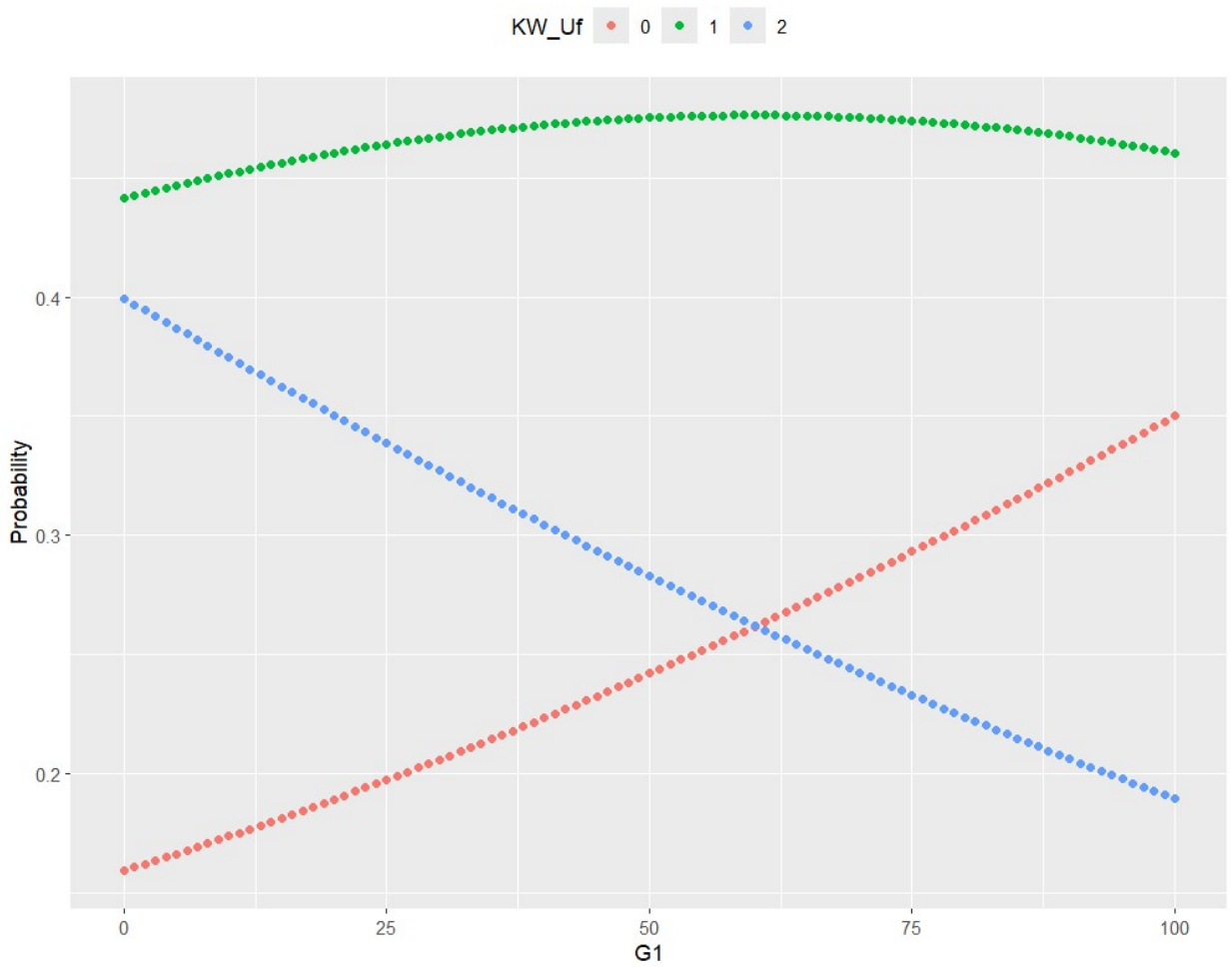

Figure S8. Probability of perceiving the "know-what" sections, KW, useful as a function of previous academic performance (G1).

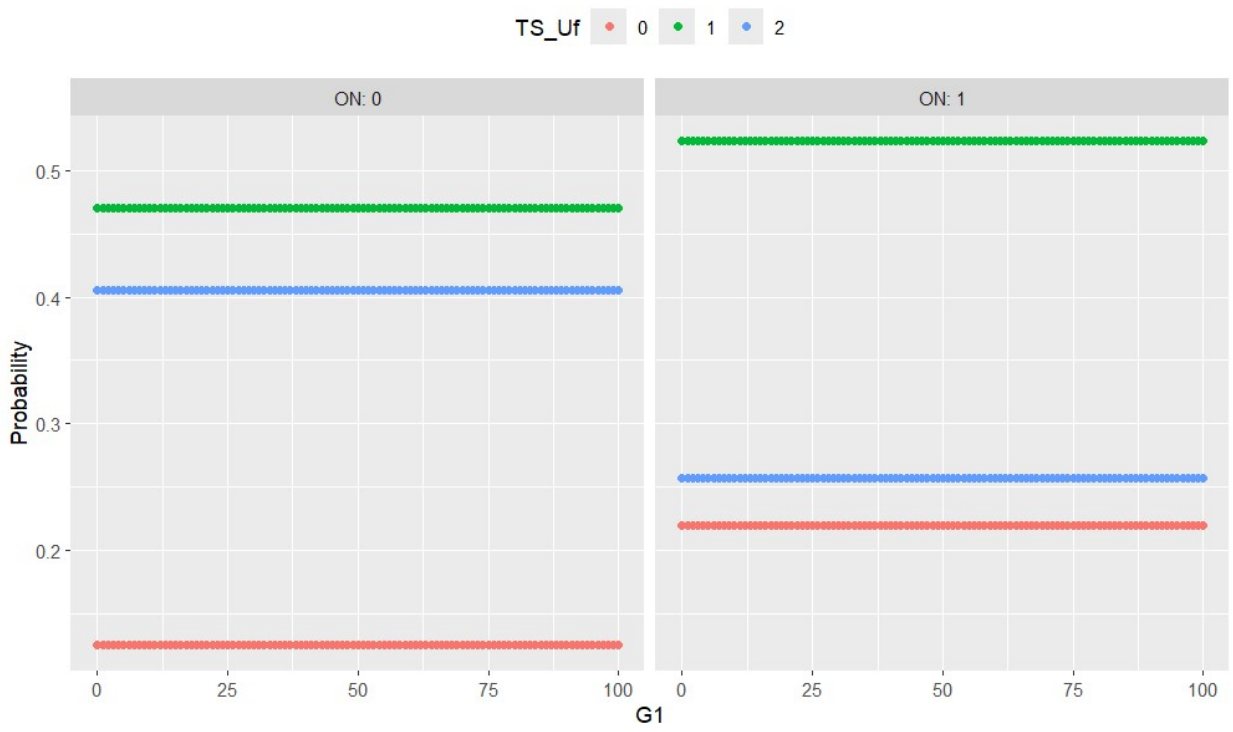

Figure S9. Probability of perceiving the tutoring sessions, TS, useful as a function of previous academic performance (G1) online coursing (ON).

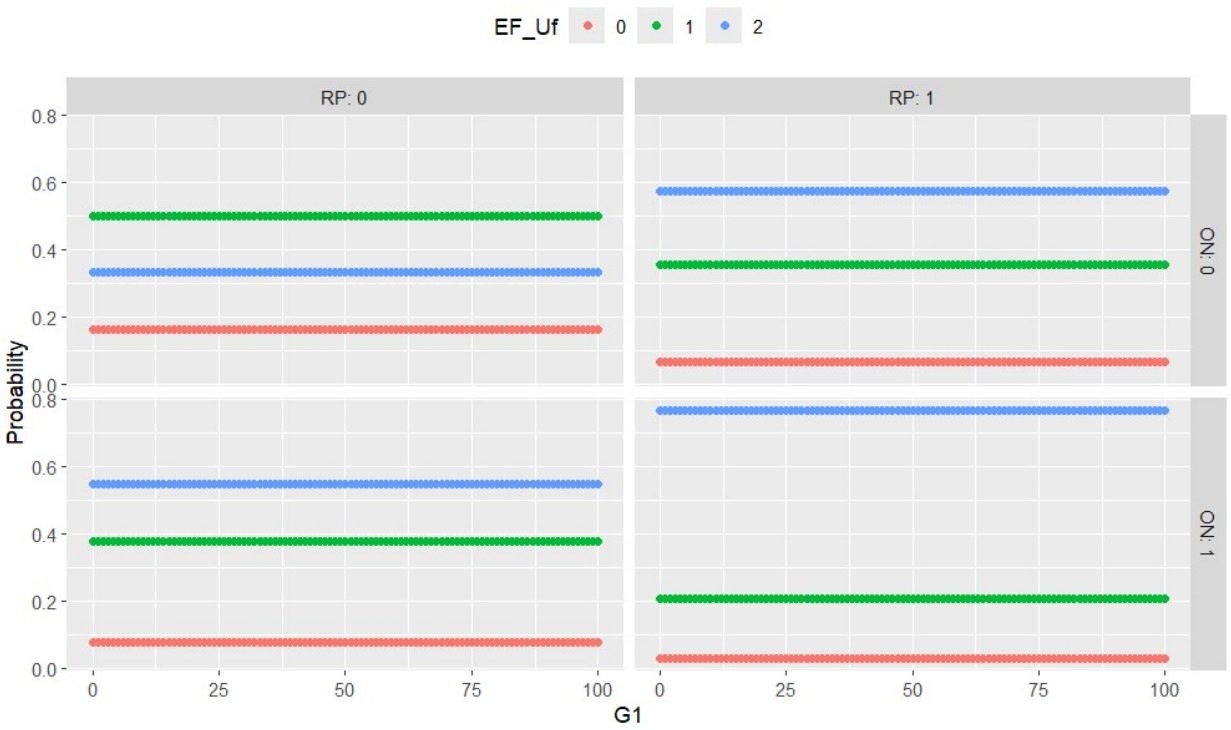

Figure S10. Probability of perceiving the electronic forum, EF, useful as a function of previous academic performance (G1), course repetition (RP) and online coursing (ON).

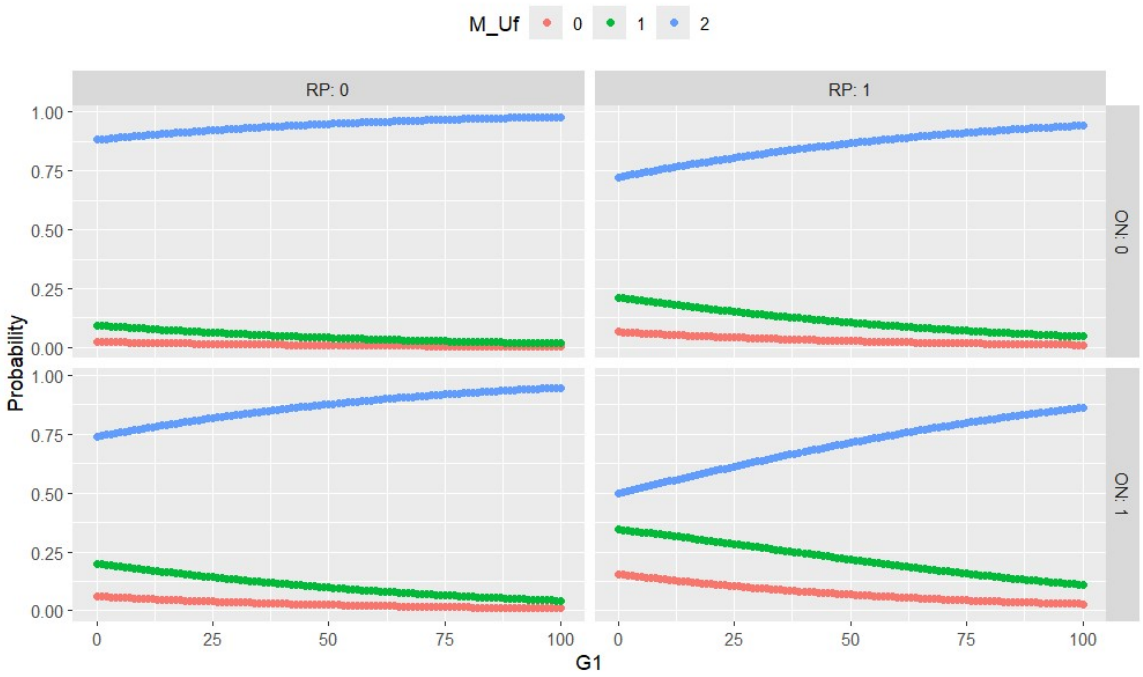

Figure S11. Probability of perceiving the mocks useful as a function of previous academic performance (G1), course repetition (RP) and online coursing (ON).

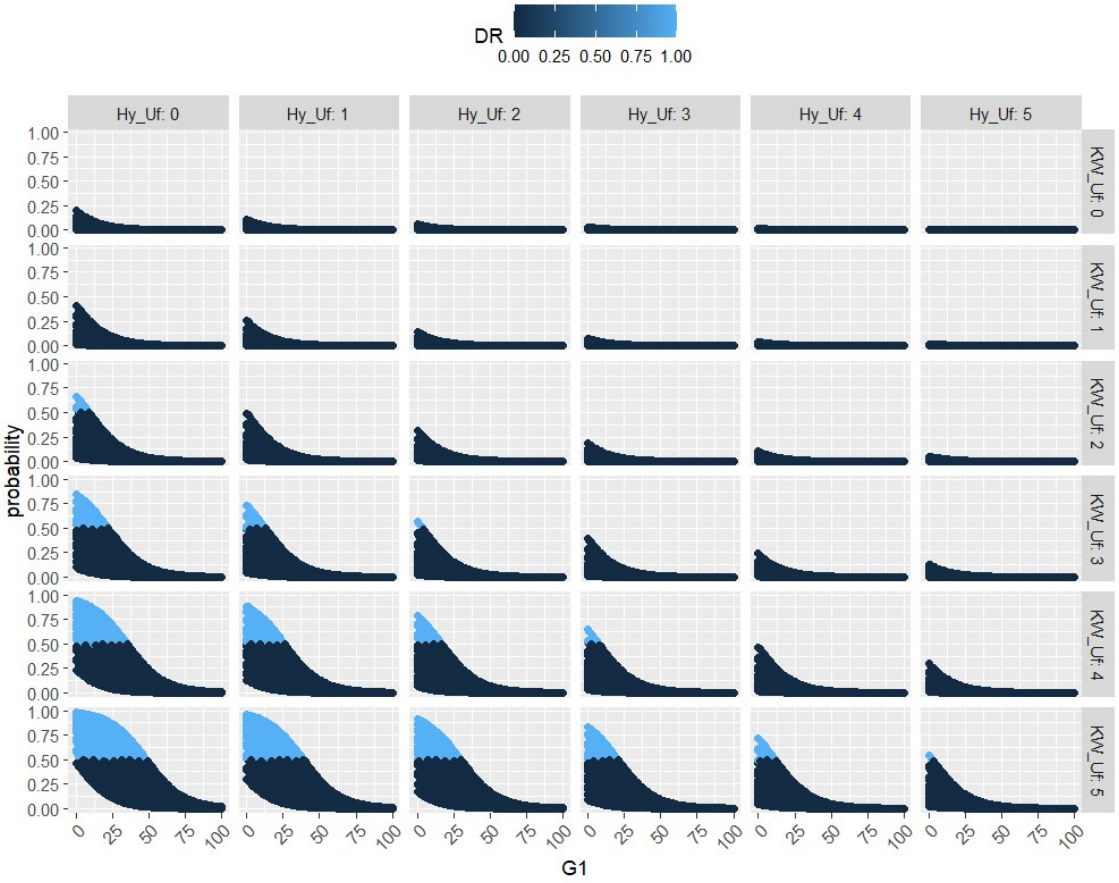

Figure S12. Probability of disengagement that leads to dropping out as a function of initial academic performance, G1, and the perceived usefulness of “know-what", KW_Uf, and “a bit of History” (Hy_Uf) sections of the interactive materials.

Supplementary sections

Section SS1. Detailed statistical discussion on students' general descriptors: the impact of the pandemic

Table S2 presents the general descriptors for students, along with their temporal evolution during the pre-pandemic, pandemic, and post-pandemic periods. Of the 4511 students included in the study, 32% identified themselves as male and 68% as female. The general dropout rate during the semester (DR) reached 19%. The overall repetition rate (RP) is 23.5%, with most of the repeater students having dropped out during the previous edition of the course. Additionally, 7.9% of students opted for the distance learning option (ON). Regarding general performance in chemistry, the overall average final grade was 29.1% (minimum grade to pass the course is 51%).

To assess the temporal evolution of these descriptors across pre-pandemic (2017–2019), early pandemic (2020), late pandemic (2021), and post-pandemic (2022–2023) periods (Table S3), statistical comparisons were performed using two-sample z-tests for binary outcomes and the Games-Howell test for the continuous outcome (GCI). Gender distribution remained stable across all periods. Pearson's chi-square test showed no significant association between gender and period: $\chi^2(1) = 0.51$, $p = 0.474$, Cramer's V = 0.015 (pre *vs*. early pandemic), $\chi^2(1) = 0.10$, $p = 0.751$, V = 0.008 (early pandemic *vs*. late pandemic), $\chi^2(1) = 0.01$, $p = 0.916$, V = 0.002 (late pandemic *vs*. post), and $\chi^2(1) = 1.44$, $p = 0.230$, V = 0.022 (pre *vs*. post).

In contrast, the proportion of repeater students significantly increased during the early pandemic. Comparing pre-pandemic to 2020, the proportion of repeaters rose by 6.1% (95% CI [1.6, 10.7]); $\chi^2(1, N = 2{,}413) = 11.24$, $p = 0.001$, V = 0.068. This rise persisted into the post-pandemic period, since there is an increase of 5.5% from pre- to post-pandemic (95% CI [1.8, 9.3]; $\chi^2(1, N = 3{,}054) = 13.43$, $p < 0.001$, V = 0.066. However, there were no significant differences between early and late pandemic ($\chi^2(1) = 0.01$, $p = 0.927$, V = 0.002) or late pandemic and post-pandemic ($\chi^2(1) = 0.16$, $p = 0.687$, V = 0.009).

On the other hand, there was an average increase of 4.9% in enrollment in the online modality during 2020 (95% CI [2.1, 7.7]; $\chi^2(1, N = 2{,}413) = 18.59$, $p < 0.001$, V = 0.088). Interestingly, despite a statistically significant 3.7% reduction in online enrollment from 2020 to 2021 (95% CI [-7.3, 0.0]; $\chi^2(1, N = 1{,}457) = 6.29$, $p = 0.012$, V = 0.066), the increased preference for this modality remained nearly intact after returning to in-person practice sessions. In fact, in the post-pandemic period, online enrollment stabilized at 4.7% above pre-pandemic levels (95% CI [2.3, 7.1]; $\chi^2(1, N = 3{,}054) = 23.47$, $p < 0.001$, V = 0.088).

The impact of the pandemic on dropout rates was immediate and strong. Between pre-pandemic and 2020, dropout increased by 19.8% (95% CI [15.4, 24.3]; $\chi^2(1) = 123.31$, $p < 0.001$, V = 0.226), indicating a medium-to-large effect size. This phenomenon occurred in conjunction with a 9.1% drop in their academic performance (95% *CI* [-11.8, -6.4]; Games–Howell $q(1268) = -12.20$, $p < 0.001$). Besides, from 2020 to 2021, dropout decreased significantly by 14.0% (95% CI [–19.8, –8.1]; $\chi^2(1) = 35.97$, $p < 0.001$, V = 0.157), and a 7.5% increase in final course marks (95% *CI* [4.3, 10.6]; $q(1439) = 8.64$, $p < 0.001$). Besides, comparing pre- and post-pandemic periods, dropout rates did not differ significantly ($\chi^2(1) = 2.15$, $p = 0.143$, V = 0.027). Regarding average final marks, GCI, even though there is a drop of 3.6% from 2017–2019 to 2022/2023 (95% *CI* [-5.8, -1.6]; $q(2920) = -6.32$, $p < 0.001$), the variation between 2019 and 2022/2023 is not statistically significant (-0.7%; 95% *CI* [-2.9, +1.5]; $q(1071) = -0.86$, $p = 0.544$). As a result, dropout rates and academic performance appear to have returned to levels comparable to those seen just before the pandemic.

Section SS2. Detailed statistical discussion on students' access to digital resources and their evolution across the pandemic

Data on the relative number of views and the total view time of recorded theory lectures (TV) and practice session videos (PV) from 2020 to 2023 were analyzed using two-factor ANOVAs (factors: year and week).

For **TV views**, results revealed a significant effect of week, $F(13, 39) = 5.01$, $p = 4.3 \times 10^{-5}$, partial $\eta^2 = 0.63$, and a significant effect of year, $F(3, 39) = 23.43$, $p = 7.7 \times 10^{-9}$, partial $\eta^2 = 0.64$. Indeed, the number of views decreased over the weeks of the course (Figures S1a and c), with an overall reduction of about 121% relative to enrollment across 2020–2023. Moreover, data from 2023 show a smaller number of views for theory lectures of about 83 - 87% of the course enrollment compared to the 2020–2022 period. In detail, post-hoc Tukey HSD tests for the Year factor showed that views in 2023 (M = 0.59) were significantly lower than in 2020 (M = 1.43, $p < 0.001$, Cohen's $d = 2.57$), 2021 (M = 1.42, $p < 0.001$, $d = 2.53$), and 2022 (M = 1.46, $p < 0.001$, $d = 2.66$). No significant differences were found between the years 2020, 2021, and 2022 (all $p > 0.98$).

For **PV views**, there was also a significant effect of week, $F(11, 33) = 11.21$, $p = 3.0 \times 10^{-8}$, partial $\eta^2 = 0.79$, and year, $F(3, 33) = 12.15$, $p = 1.6 \times 10^{-5}$, partial $\eta^2 = 0.52$. View numbers declined across weeks, with a reduction of 71% of the course enrollment for the whole sample 2020–2023. Besides, recorded practice sessions from 2022–2023 also have fewer views compared to the pandemic period, with a reduction between 29 and 43% of the enrollment. In detail, post-hoc Tukey HSD tests for the Year factor indicated that views in 2022 (M = 0.81) and 2023 (M = 0.72) were both significantly lower than in 2020 (M = 1.10; 2022 *vs*. 2020: $p = 0.010$, $d = 1.37$; 2023 *vs*. 2020: $p < 0.001$, $d = 1.79$) and 2021 (M = 1.15; 2022 *vs*. 2021: $p = 0.002$, $d = 1.63$; 2023 *vs*. 2021: $p < 0.001$, $d = 2.05$). Views in 2022 and 2023 did not differ significantly from each other ($p = 0.73$).

For **TV view time**, there was a significant effect of week, $F(13, 39) = 3.00$, $p = 0.004$, partial $\eta^2 = 0.50$, and year, $F(3, 39) = 58.59$, $p = 1.6 \times 10^{-14}$, partial $\eta^2 = 0.82$. View times were higher in 2020 compared to 2021–2023, with differences corresponding to 11.1% to 13.5% of the video duration. Post-hoc Tukey HSD tests for the Year factor revealed that view time was significantly longer in 2020 (M = 35.44%) compared to 2021 (M = 24.34%, $p < 0.001$, $d = 3.60$), 2022 (M = 21.96%, $p < 0.001$, $d = 4.38$), and 2023 (M = 22.70%, $p < 0.001$, $d = 4.14$). There were no significant differences in view time among the years 2021, 2022, and 2023 (all $p > 0.18$).

For **PV view time**, a significant effect of week emerged, $F(11, 33) = 12.68$, $p = 6.6 \times 10^{-9}$, partial $\eta^2 = 0.81$, as well as a significant effect of year, $F(3, 33) = 43.75$, $p = 1.3 \times 10^{-11}$, partial $\eta^2 = 0.80$. View times were highest in 2020, with engagement between 4–9% longer than in subsequent years. In detail, post-hoc Tukey HSD tests for the Year factor showed that view time in 2020 (M = 32.33%) was significantly higher than in all subsequent years (2021: M = 28.49%, $p < 0.001$, $d = 1.98$; 2022: M = 23.35%, $p < 0.001$, $d = 4.63$; 2023: M = 27.15%, $p < 0.001$, $d = 2.67$). Furthermore, view time in 2022 was significantly lower than in 2023 ($p < 0.001$, $d = 1.96$). No other comparisons were significant.

Section SS3. Statistical detailed discussion on survival analysis to analyze dropout

For each student enrolled in the 2022 in-person course (571 students), the number of weeks they attended in-person practice sessions was collected. Disengagement point was defined as the week in which a student stopped attending practice sessions consecutively until the final week of classes (week 12; see the statistical analysis section for further details). It can be observed that disengagement probability increases notably during the second half of the course (Figure 5a; week 6 onwards). In general, assuming the sample is representative of the population, it is predicted that 9.7% of students drop out in the second half of the course (95% CI [-11.0, -8.4]), compared to a three times lower dropout rate in the first half (3.1%; 95% CI [1.7, 4.5]).

A Cox Proportional Hazards Model was applied to the data to identify statistically significant dropout determinants influencing the risk of dropout. The results are summarized in Table 3. The model's concordance index is 76%, indicating good predictive power in distinguishing between students who drop out and those who remain enrolled. Consistent with the logistic regression findings, higher test scores are associated with a lower probability of dropping out. Specifically, for each additional point in the first test, G1, the hazard of dropping out decreases by approximately 6% ($p = 7 \times 10^{-8}$). Interestingly, repeaters are 72% more likely to drop out than freshman students when all other variables are held constant ($p = 0.02$). The disparity in dropout rates between repeater and freshman students becomes more pronounced in the second half of the semester.