This article is published by Taylor & Francis in the European Journal of Engineering Education.

Cover page

Title: Engineering students' approaches to learning in two student-centred instructional models before and during the COVID-19 pandemic

Authors: Johanna Rämö[1], Petri Nokelainen[2], Riikka Kangaslampi[1], Elina Viro[1], Terhi Kaarakka[1], Maiju Nieminen[1], Jani Hirvonen[1], Simo Ali-Löytty[1]

Affliations:
[1] Tampere University, Faculty of Information Technology and Communication Sciences, Tampere, Finland
[2] Tampere University, Faculty of Education and Culture, Tampere, Finland

Corresponding author: Johanna Rämö, johanna.ramo@uef.fi

ORCID:
Johanna Rämö: orcid.org/0000-0002-9836-1896
Riikka Kangaslampi: orcid.org/0000-0002-9059-8047
Petri Nokelainen: orcid.org/0000-0002-8195-7001
Elina Viro: orcid.org/0000-0001-6644-4251
Terhi Kaarakka: orcid.org/0000-0002-3824-8939
Jani Hirvonen: orcid.org/0000-0003-1667-2760
Simo Ali-Löytty: orcid.org/0000-0002-6720-7722

Funding details: Not applicable.

Disclosure statement: There is no financial interest or benefit that has arisen from the direct application of our research.

Data availability: Not applicable

Ethics Approval: Ethical committee of Tampere University has given permission for this research.

Acknowledgements:
The authors would like to thank all other teachers of the courses considered in this study, especially Jussi Kangas and Petteri Laakkonen.

Biographical note:
Johanna Rämö is a university lecturer at the Computing Sciences unit, Tampere University, Finland. She received a Ph.D. in mathematics from Queen Mary, University of London, United

Kingdom. Her research concerns undergraduate mathematics education, focusing on learning environments and novel assessment methods. Footnote: Johanna Rämö is currently a university lecturer at the Department of Physics and Mathematics in the University of Eastern Finland.

Petri Nokelainen is a full professor of Engineering Pedagogy in Tampere University, Finland. His current research interests include motivation, self-regulation, emotions and basic psychological needs satisfaction in the contexts of engineering higher education and working life.

Riikka Kangaslampi is working as a university lecturer in mathematics at the Computing Sciences unit, Tampere University, Finland. She received her degree of Doctor of Science in Technology from Helsinki University of Technology, Finland, in 2008. Her main research interests are discrete geometry and university mathematics education.

Elina Viro is a university instructor in mathematics at Tampere University and is writing her doctoral thesis on project-based learning in mathematics at the lower secondary school level. She has coached first-year engineering mathematics students. In addition, she has conducted pedagogical courses for mathematics pre-service teachers. She is interested in developing student-centered teaching and learning.

Terhi Kaarakka is a university lecturer in mathematics at Tampere University. Terhi has made a doctoral dissertation on stochastics. She teaches engineering mathematics for large groups and therefore she is interested in how technology can support mathematics learning

Maiju Nieminen is a secondary school teacher of mathematical sciences. She received her degree of Master of Science (Technology) from Tampere University in 2021.

Jani Hirvonen is a university instructor in mathematics at the Computing Sciences unit, Tampere University. While his doctoral dissertation handled theoretical mathematics, he's nowadays interested in university mathematics education in general and especially with the support of technological means.

Simo Ali-Löytty is working as a university lecturer in mathematics at the Computing Sciences unit, Tampere University, Finland. He has long pedagogical and technical experience in using and designing digital learning activities. His current research interest is technology-supported mathematics education.

Engineering students’ approaches to learning in two student-centred instructional models before and during the COVID-19 pandemic

Abstract

In this study, we investigated how well two different student-centred instructional models fostered engineering students’ learning in a time of crisis. We analysed students’ (N = 375) approaches to learning during four engineering mathematics courses in a Finnish university before and during the COVID-19 pandemic. Students’ deep and surface approaches to learning, as well as organised studying, were measured five times during an eight-month period. For the control group the student-centred elements were added to the framework of traditional lecture-based teaching, whereas the intervention group’s instructional model disrupted the structures of traditional teaching more profoundly. Our results indicate that the pandemic and related restrictions were linked to a decrease in students’ deep approach to learning and organised studying, and an increase of surface approach to learning in both groups. However, the intervention group’s instructional model supported the deep approach to learning better than that of the control group.

Keywords: approaches to learning; student-centred teaching; flipped learning; pandemic; engineering mathematics

Introduction

With the emergence of the COVID-19 pandemic in 2020, people worldwide faced sudden changes in their lives and were forced to adjust to new, stressful situations. Social contacts were restricted drastically, and people needed to work and study from their homes online. University students were stressed over their emotional health, academic work, employment, finances and their loved ones’ health (Mushquash & Grassia, 2020). The pandemic negatively affected students’ well-being and fuelled anxiety, causing depression symptoms (Evans et al., 2021). For these reasons, a need exists to find instructional models that support students' well-being and keep the quality of learning high, even in exceptional circumstances.

One strategy for supporting the quality of students’ learning are student-centred teaching methods. Student-centred teaching can be seen as an umbrella that encompasses teaching methods in which students have an active role in their knowledge construction (Baeten et al., 2013). It has been shown to forster skills that are important for students’ future careers – such as problem solving, communication and teamwork (e.g., Baytiyeh & Naja, 2017; Yadav et al., 2011) – and it can affect students’ learning positively (e.g., Baytiyeh & Naja, 2017; Dochy et al., 2003; Lahdenperä et al. 2018; Laursen et al., 2014; Yadav et al., 2011). One could hypothesise that student-centred teaching supports the quality of students’ learning also during the pandemic.

In this study, we investigated how two instructional models that applied student-centred teaching in different intensities, supported the quality of learning before and during the pandemic. We viewed the quality of learning from the perspective of students’ approaches to learning (SAL; Marton & Säljö, 1976; Parpala & Lindblom-Ylänne, 2012). Approaches to learning describe intentions that students have when approaching a learning situation and typically are divided into deep and surface approaches, together with organised studying. Studies concerning the relationship of student-centred teaching methods and students' approaches to learning report mixed results. Many studies have found that student-centred learning environments support favourable approaches to learning (e.g., Dolmans 2016; Wilson & Fowler, 2005), while in other studies student-centred learning environments have been linked with less optimal approaches to learning (e.g., Baeten et al., 2013; Struyven et al., 2006). This implies that more information is needed on different kinds of student-centred learning environments and contexts in which they are applied when studying their effect on students' approaches to learning.

Student-centred teaching methods

During the past few decades, student-centred teaching methods have gained wide interest in higher education (e.g., Baeten et al., 2010; Karabulut-Ilgu et al., 2018). Baeten et al. (2013) define student-centred teaching as having three characteristics: Students play an active role in constructing knowledge; teachers facilitate students' work; and the assignments used in teaching are authentic. Student-centred teaching can take different forms, and many teaching models can be considered to fall under its umbrella, including active learning (e.g., Prince, 2004), problem-based learning (e.g., Dochy et al., 2003), inquiry-based learning (e.g., Laursen et al., 2014) and flipped learning (Talbert, 2017).

In the context of engineering education, student-centred teaching methods have been used and studied widely. In their systematic review Karabulut-Ilgu et al. (2018) investigated 62 studies on flipped learning in engineering education. According to their review, the learning gains in flipped learning were as good as or better than in traditional lecture-based teaching. As benefits of flipped learning, they reported flexibility of teaching arrangements, student engagement, learning gains in transferable skills, and improvement in interaction both between students and teachers and among students. According to the review, challenges include increased workload for teachers, student dissatisfaction with the teaching materials, technological issues and student resistance. In their three-year follow-up study, Polanco et al. (2004) showed that problem-based learning improved engineering students' academic achievement more than traditional teaching. The positive effect of problem-based learning was seen also in consecutive courses. The results of Yadav et al. (2011) suggest that problem-based learning fosters problem-solving skills better than a lecture-based method for engineering students.

Students' approaches to learning

There has been a long tradition of studying students' approaches to learning in higher education research. Its origins lie in a study by Marton and Säljö (1976). They distinguished between two ways of processing information: surface processing and deep processing. Later, these concepts were developed further and replaced by *surface approach to learning* and *deep approach to learning*, which include students' intentions related to both their studying and learning processes (e.g., Entwistle & Ramsden, 1983; Entwistle et al., 2006). Students applying the deep approach to learning aim to understand ideas for themselves, relate ideas to their previous knowledge and examine arguments critically. Students who apply the surface approach to learning use unreflective strategies such as memorisation and carrying out procedures (Entwistle & Peterson, 2004). They study in an unreflective manner and their knowledge is fragmented (Lindblom-Ylänne et al., 2019). There is also a third approach, which has been referred to as strategic approach (Entwistle & Ramsden, 1983) or achieving approach (Biggs, 1993). Nowadays it is usually called *organised studying*. Students who apply this approach study in an organised manner, are good at time management, and can manage their concentration and effort (Entwistle & McCune, 2004).

The deep approach to learning has been linked to higher achievement than the surface approach to learning (e.g., Marton & Säljö, 1976; Minbashian et al., 2004). Also organised studying has been shown to correlate positively with study success (Asikainen et al., 2014; Rytkönen et al., 2012). In particular, the combination of deep approach to learning and organised studying has been noted to be favorable (Haarala-Muhonen et al., 2017).

Several characteristics of students and learning environments have been linked to students' approaches to learning. If a student finds that a course's workload is too heavy or experiences stress, it can manifest as the student taking the surface approach to learning (Cheung et al., 2020). Also, a lack of challenges can lead to the same result (Coertjens et al., 2016). The deep approach to learning and organised studying have been associated positively with the perception of receiving peer support (Coertjens et al., 2016; Lahdenperä et al., 2018). A study by Du et al. (2019) conducted with engineering students suggested that team projects foster a deep approach to learning.

Student-centred learning environments can foster the deep approach to learning (e.g., Dolmans et al., 2016; Wilson & Fowler, 2005), but they can also encourage students to use the surface approach to learning (e.g., Baeten et al., 2013; Struyven et al., 2006). Leung et al. (2008) found that for engineering students, teacher-centred teaching correlated with the surface approach to learning at Hong Kong universities and with the deep approach to learning in mainland China. Therefore, comparing different student-centred learning environments is of interest in order to find out which elements of learning environments foster favourable approaches to learning.

In this study, we investigated engineering students' approaches to learning in two different student-centred instructional models which utilised student-centred teaching in different intensities. One of the student-centred instructional models functioned within the traditional framework of lectures, and the other disrupted the traditional structures of mathematics teaching more profoundly. Even though the control and intervention groups were exposed to different kinds of teaching methods, both groups received student-centred instruction from pedagogically qualified and motivated teachers. We took a longitudinal approach by measuring students' approaches to learning in four consecutive courses, as short interventions might not be sufficient to impact students' approaches to learning (Baeten et al., 2010; Karabulut-Ilgu et al., 2018; Wilson & Fowler, 2005). A novel perspective is to study the effect from a sudden crisis on students' approaches to learning. During the fourth course, the COVID-19 pandemic elicited severe disruptions in teaching, affecting both students and their learning environments drastically. All teaching had to migrate onto online platforms in mere days, and students and teachers' personal lives were restricted severely. We investigated how the two different instructional models supported student learning in this unexpected and stressful situation.

Research questions:

RQ1: Did students' approaches to learning differ over time in engineering mathematics courses that were taught using the two different instructional models?

RQ 2: Was there a change in students' approaches to learning in engineering mathematics courses taught using the two different instructional models before and during the COVID-19 pandemic?

Context

This study's participants were students taking compulsory first-year engineering mathematics courses in a research-intensive university in Finland. Altogether, four focus courses were used in this study, each worth five ECTS (European Credit Transfer and Accumulation System) credits. Course topics included functions, complex numbers, matrices, differential and integral calculus, probability and statistics. Each course lasted seven weeks, plus an exam week.

The participants were divided into two groups based on the instructional model used. The control group was taught using a student-centred model that was built within a traditional lecture-based format. The intervention group's courses were taught using a student-centred model in which the structures of traditional teaching were removed. The courses implemented with the two models had the same learning objectives and utilised the same written course materials and educational videos. In both models, students completed weekly tasks, some of which were online tasks with automatic feedback and others pen-and-paper tasks. Both models utilised the online platform Moodle. Students' grades were based on tasks completed during the course plus a final exam. The two instructional models differed mainly in how contact teaching and assessment were organised. These differences are described below and summarised in Table 1.

The control group's model comprised a weekly schedule of four hours of lectures and two two-hour exercise sessions. The lectures were sessions of approximately 250 students. In the lectures, the teacher explained and motivated the topics of the week and broke down thinking behind mathematical proofs. The students also discussed examples given by the teacher in small groups. Students completed tasks each week. During the exercise sessions, a group of approximately 25 students solved tasks in smaller groups and discussed worked-out solutions with a teaching assistant. A tutoring lab was available for studying new topics and solving tasks with help from teaching assistants. Students also could participate in a weekly basic skills support session aimed at students who need to brush up on prior knowledge.

In the intervention group, the idea of flipped learning (Talbert, 2017) was utilised. The students were introduced to new concepts with structured activities, and the group learning space was dedicated to interaction. Every week, the students received a theory pack comprising references to the written course materials and educational videos, as well as a problem set. Contact teaching comprised an exercise session, tutoring lab and basic skills support session just as for the control group. At the end of the week, students participated in a *prime-time* session (see Koskinen et al., 2018) with their teacher. The students worked in dedicated small groups of approximately eight people both in the exercise sessions and prime-time meetings. The prime-time meetings were sessions of approximately 25 students, in which each group had a 30-minute conversation with the teacher. The teacher and students discussed topics that were unclear after self-studying as well as study skills. When a student group was not discussing with a teacher, they worked on group tasks that summed up or expanded the week's topics. Similar to the control group, the intervention group was given weekly, individual tasks. The main difference was that some of the intervention group's tasks aimed at developing conceptual understanding. Also, some of the tasks were self- and peer-assessed. In addition to self-assessing their solutions to tasks, students regularly self-assessed their competencies using the course's learning objectives.

The COVID-19 pandemic closed the society on March 19, 2020. Schools at all levels were closed for face-to-face education and people's options to go out were very limited. In our focus university the teaching arrangement changed abruptly. For the control group, lectures were replaced by short educational videos. Instead of the exercise sessions, the students submitted all their tasks through the course's online platform. Solutions to tasks were self-assessed or peer assessed. Support for exercises took place in a virtual tutoring lab using Microsoft Teams. The control group's exam had two parts. The first part was an open-book online exam, and the second part comprised slightly randomised questions. The exam did not have overseers, and all course materials could be used.

In the intervention group, the exercise sessions were cancelled. The tasks that previously were done before the exercise sessions were now self-assessed by the students. The tutoring lab moved from a face-to-face setting to online teaching, similar to the control group. Prime-time sessions were held with video conferencing separately for each small group. The exam comprised an open-book online exam, an individual assignment and a brief one-on-one assessment discussion with the teacher via a video conference.

During the four courses, the control group had two different responsible teachers. On average, they had 18 years of teaching experience (min. 16, max. 20 years). For the intervention group, there were four different responsible teachers with 16 years of experience on average (min. 6, max. 23 years). All the responsible teachers had a 60-ECTS-credit pedagogical qualification. In addition, there were several teaching assistants supervising the exercise sessions and teaching in the tutoring lab. The teaching resources allocated for the control and intervention groups did not differ to a considerable degree.

Table 1. Teaching arrangements for two student groups, control and intervention, that were taught using two different student-centred instructional models.

[INSERT TABLE 1 HERE.]

Method

Participants

The sample comprised 374 first year engineering students in higher education (126 females, 33.7%; one person's gender information was unavailable). Their average age was 21.91 (*SD* = 2.336) years. Age information for three persons were missing. The females' average age (*M* = 21.98, *SD* = 2.679) was close to that of the males (*M* = 21.89, *SD* = 2.145).

The students were assigned to either the control (*n* = 200, 53.5%; 72 females, 36.0%) or the intervention group (*n* = 174, 46.5%; 54 females, 31.0%) according to their study programmes.

The students in the control group were from the electrical engineering, bioengineering and information technology programmes, whereas the students in the intervention group were from automation engineering, mechanical engineering, materials science and environmental and energy engineering. Gender distribution in the two groups was close to the overall gender distribution of the sample. The average ages in the two groups were similar and close to the whole sample average age: Control M = 22.00 (SD = 2.443), Intervention M = 21.81 (SD = 2.210).

Even though the students in the control and intervention groups had different majors, they all were engineering students whose first-year studies do not differ much from each other. Also, an initial measurement was done in the beginning of the semester (timepoint t0) to verify that students in the two groups did not have different approaches to learning.

Procedure

Students' self-assessments of their approaches to learning were collected using an online survey on the Moodle platform used for teaching. Consent to participate was collected from the students during the first online survey data collection. Participation in this study was voluntary and the students were aware that they could withdraw at any time or refuse to answer any question without any consequences. The students completed five surveys during four engineering mathematics courses (each course lasted about seven weeks). The intervals between the five measurements varied from five to ten weeks. The first measurement (t0) took place in August 2019, before the first course. The next three measurements were taken at the beginning of the second (t1, October 2019), third (t2, January 2020) and fourth (t3, March 2020) course. The fifth measurement (t4) was taken at the end of the fourth course (April 2020) during the pandemic.

Instrument

The participants completed a questionnaire (Parpala & Lindblom-Ylänne, 2012) on approaches to learning five times during this study. The questionnaire was based on the Experiences of Teaching and Learning Questionnaire (ETLQ; Entwistle et al., 2003) and evaluated university students' approaches to learning in a state/event (course-specific) level with 12 self-response items under three dimensions: deep approach (four items), surface approach (four items) and organised studying (four items). Responses to the items were provided on a five-point scale ranging from 1 (totally disagree) to 5 (totally agree). The original study used to validate the questionnaire (Parpala, 2010) employed two university student samples from Finland and the UK and showed the following internal consistency values: deep (α_{Fin} = 0.82, α_{Brit} = 0.76); surface (α_{Fin} = 0.59, α_{Brit} = 0.70); organised (α_{Fin} = 0.76, α_{Brit} = 0.78). A recent study with Danish university students (Herrmann et al., 2017) reported internal consistency values of 0.73 (deep), 0.77 (surface) and 0.77 (organised). In our study, the average internal consistency values of the five repeated measurements were 0.52 (deep), 0.67 (surface) and 0.79 (organised).

Statistical analyses

Multilevel modeling (MLM, see e.g., Finch et al., 2014; Hox, 2010), more specifically mixed effects growth curve modeling, was performed with R statistical computing environment (R Core Team, 2020; RStudioTeam, 2016) to investigate change over time. Three SAL dimensions (deep approach to learning, surface approach to learning, organised studying) were the dependent variables (DV) in the analysis while Time, Treatment (0 = control, 1 = intervention) and Gender

(0 = female, 1 = male) were the predictors. Predictors were not centered by the grand mean or group mean as they contained interpretable zero values. MLM was used as it allows analysis between different levels (level 1: SAL dimensions, level 2: students) and inclusion of variables with missing data. All the analyses were conducted with both *nlme::lme* (Pinheiro et al., 2021) and *lme4::lmer* (Bates et al., 2015) programs and maximum likelihood (ML) and restricted maximum likelihood (REML) methods. Although it is known that ML estimates of variances are underestimated (e.g., Hoffman, 2014), the effect is small with large data.

Before analysis, the data were analysed for missing values with *mice* (van Buuren & Groothuis-Oudshoorn, 2011), *finalfit* (Harrison et al., 2021) and *naniar* (Tierney et al., 2021) *R* packages. Original data contained five repeated measurements (total number of 2025 observations) from 405 students with 25.5% of missing data. Missing data analysis (e.g., Gelman & Hill, 2007; van Buuren, 2018) with *mice::md.pattern* showed that the most frequent pattern (1508 observations, 74.5%) was the one that contained complete data. The second most frequent pattern (508 observations, 25.1%) had missing data in the survey variables (Deep, Surface, Organised).

Investigation of missing data across the five timepoints showed that the proportion of missing data had a general declining tendency over time in SAL (30.9%, 31.4%, 20.2%, 21.7%, 21.2%) variables. Missing data were present quite equally in both control (27.3%) and intervention (22.5%) groups. Investigation of missing data matrix (*finalfit::missing_pairs*) showed that the most frequent missing values (in SAL variables) were randomly distributed across other variables. In addition, missing data was visualized with *naniar::geom_miss_point* by replacing missing values with values 10% lower than the minimum value of that variable (inspired by the 20% rule from Swayne & Buja, 1998). Inspection of missing data through the five timepoints showed that it was evenly distributed in both control and intervention groups. As the data contained quite a large number of complete observations (1508 out of 2025), we decided to include observations containing missing values in the analysis and continue the preliminary analysis with the original data (2025 observations from 405 students).

After missing value analysis, the data were investigated by each DV (Deep, Surface, Organised) for assumptions related to 1) within-group error and 2) random effects (Pinheiro & Bates, 2000, pp. 174-196). Analyses of the within-group residuals (BLUP's, best linear unbiased predictors of the within-group errors) with both homoskedastic and heteroskedastic (different variances for each level of Treatment) models showed that the errors were centered at zero and normally distributed. Plotting the observed responses versus the within-group fitted values for each DV showed small standardised residuals, suggesting that the linear mixed-effects model was successful in explaining the growth curves. Normal plots of estimated random effects from both homoskedastic and heteroskedastic models for each DV showed symmetrical distributions for the intercept and time. Scatter plot matrixes of the estimated random effects indicated that there were no departures from the assumption of homogeneity of the random-effects distribution. Extreme values for each DV (using Time, Treatment and Gender as predictors) were analysed with *RModelDiagnostics* package (Wiley, 2020). Results indicated that there were 31 students who had extreme values in the DV's. After excluding these students from the analysis, multivariate normality was achieved. The final data used in the analysis contained 1379 observations from 374 students. It retained the original gender/treatment distributions and average age of the participants.

Balance of the control and intervention groups was studied with full propensity score matching (PSM) using the *MatchIt* package (Ho et al., 2011) without replacement. The propensity score was estimated using logistic regression of the Treatment on the covariates (Deep, Surface, Organised). Results indicated that the mean variance ratio (preferably close to 1) and the mean empirical cumulative density function (preferably close to 0) values for the full data ($M_{VarRatio}$ = 0.82; $SD_{VarRatio}$ = 0.048; $Min_{VarRatio}$ = 0.79; $Max_{VarRatio}$ = 0.89) were quite close to the values of the matched data ($M_{VarRatio}$ = 0.96; $SD_{VarRatio}$ = 0.109; $Min_{VarRatio}$ = 0.80; $Max_{VarRatio}$ = 1.04). As the difference between the unmatched and matched data was small, further analyses were conducted without matching.

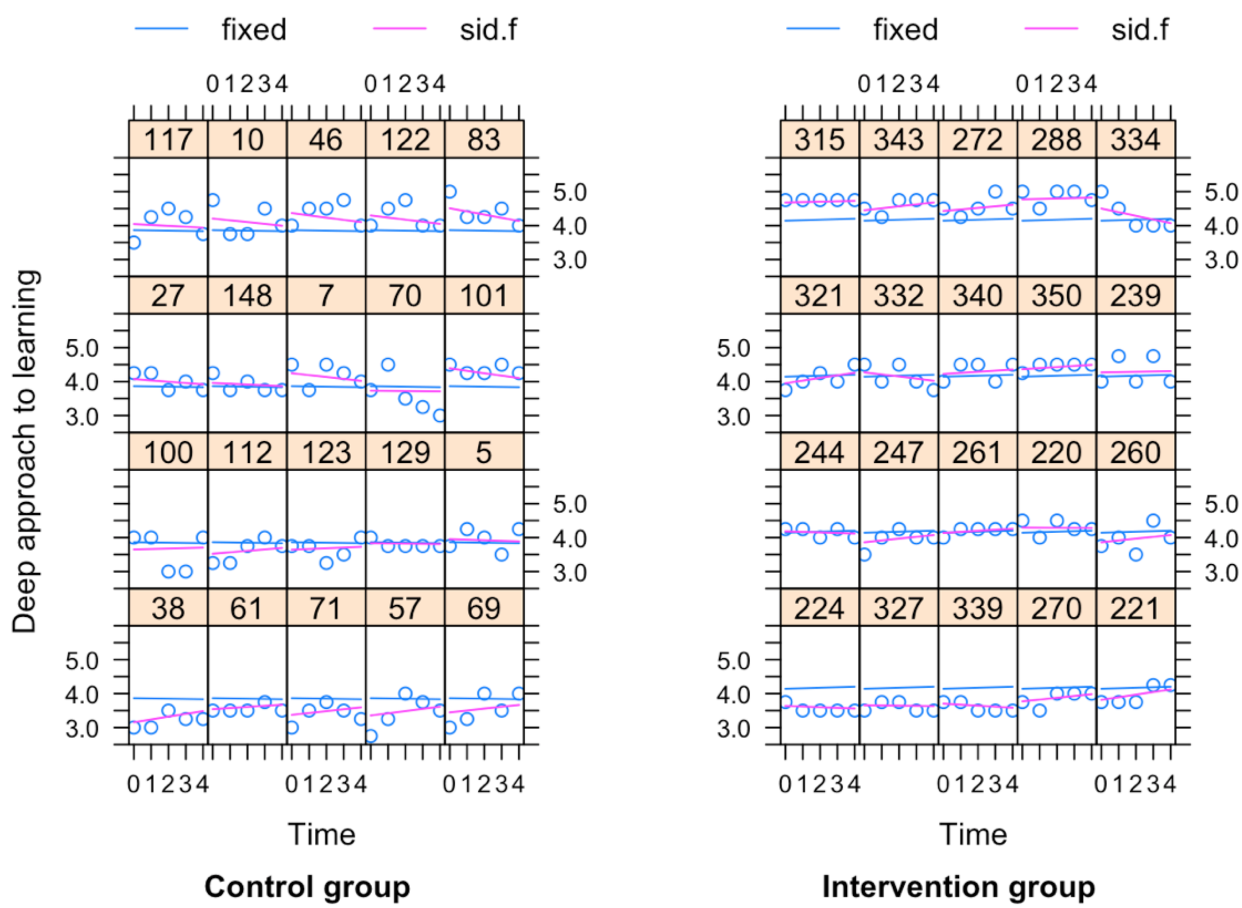
The modeling strategy in this study was traditional (e.g., Pinheiro & Bates, 2000), comparing an unconditional means model (a.k.a. "empty", "baseline" or "null" model) without predictors to several conditional models with an incremental number of predictors (including Time, Treatment and Gender). To estimate the general developmental trends for students' approaches to learning, a two-level growth model of longitudinal change (Raudenbush & Bryk, 1986) was applied to the data. The first level (within-student) model describes each student's intercepts (initial status at each measurement point) and slopes (rate of linear growth). The second level (between-student) model describes individual differences of students in these growth curve parameters. Within and between student variability were investigated with three unconditional means models (Model 0: random intercepts and no predictors; Model 1: random intercepts and time as a predictor; Model 2: random intercepts and slopes, time as a predictor) to see if there was a systematic mean-level change and individual variability in the approaches to learning and studying.

Before proceeding with the investigation of the properties of these three models, intraclass correlation (ICC) values were calculated with the *performance::icc* package (Lüdecke et al., 2021) to determine the possible need for multilevel modeling. The ICC value describes the variance explained in a DV due to between-person clustering in the data. A high ICC value indicates that a multilevel modeling approach is relevant as a relatively high proportion of the variance in the DV is explained by the clustering structure in the data. Results showed quite high ICC values (Deep: Model 0_{ICC} = 0.55, Model 1_{ICC} = 0.55, Model 2_{ICC} = 0.60; Surface: Model 0_{ICC} = 0.59, Model 1_{ICC} = 0.61, Model 2_{ICC} = 0.69; Organised: Model 0_{ICC} = 0.67, Model 1_{ICC} = 0.68, Model 2_{ICC} = 0.72), indicating a clear need for the use of a multilevel modeling approach with all three DV's.

Need for random intercepts (Model 1) or random intercepts and slopes (Model 2) were investigated by comparing the two models by their AIC (Akaike Information Criterion) values and with likelihood ratio (χ^2) test (*car::anova*, Fox & Weisberg, 2019). These analyses were conducted with all three DV's (Deep, Surface, Organised), but for illustrative purposes we present here only results related to deep approach to learning. AIC values for the two models were 1462.348 for Model 1 and 1443.427 for Model 2. The Model 2 with random intercepts and slopes had a better fit (lower AIC value) than Model 1. Subtracting the AIC value for Model 1 from the (AIC_{min}) value for Model 2 produced a difference larger than 2, indication of less than substantial evidence that Model 1 is superior compared to Model 2 (Burnham & Anderson, 2002, p.70). The same result was produced by likelihood ratio test (LRT): Model 2 ($log.Lik$ = -715.713) outperformed Model 1 ($log.Lik$ = -727.174), p < .001). Similar results were obtained for the other two SAL dimensions. Based on these findings, analysis was continued with Model 2.

As Figure 1 with 20 randomly selected students from the control and intervention groups illustrates, model fit was improved by allowing both individual variation of students' average scores on the deep approach to learning and individual variation of the related slopes.

Figure 1. Random intercepts and slopes model of deep approach over time with randomly selected students from control (n = 20) and intervention (n = 20) groups



Note: Blue line ("fixed") represents overall estimated trend and red lines ("sid.f") represent students' individual estimated trends.

The next step was to investigate if the cluster level error structures (five repeated measurements clustered by students) were autocorrelated. Each of the aforementioned models were re-run in *nlme::lme* with *corCAR1* function (instead of *corAR1*) as it allows for unequally spaced time between observations (Finch, Bolin, & Kelley, 2014).

In this study, timepoints were unequally spaced as the distance between measurements varied from five to ten weeks, and not all students completed the five surveys. AIC values for the autocorrelation error structure models for Deep and Surface dimensions were slightly better (smaller) than for the previously presented Model 2 (assuming the independence of random effects): $AIC_{Model2Deep}$ = 1443.427; $AIC_{Model2DeepCAR1}$ = 1442.288; $AIC_{Model2Surface}$ = 2160.864; $AIC_{Model2SurfaceCAR1}$ = 2160.178. However, the AIC value differences were less than two units (Burnham & Anderson, 2002) and the BIC values were better (smaller) in all models with independent error structures. In addition, as the likelihood ratio tests were not statistically significant between independent and autocorrelated models, further analyses with the three SAL dimensions were based on an assumption of independent cluster level error structures.

Finally, as the goal was to investigate the change over time in three different DV's, level 1 orthogonal polynomial effects (e.g., Hoffman, 2015) were investigated with the Model 2 (Time as the predictor, random intercepts and slopes). Although curvilinear (cubic) trends were present in Surface and Organised models, LRT showed no significant improvement in the model fit. This led us to use a linear growth model with independent cluster level error structures for all three SAL dimensions.

To sum up, unconditional growth model with linear time as a predictor (random intercepts and slopes) will be used later as the baseline model against which other models will be compared.

Results

The descriptive statistics are presented in Table 2. The means across the five measurement points indicated a general decline in deep approach to learning and organised studying. The decline in the control group seemed a little steeper in these dimensions than in the intervention group. Although the trend for surface approach to learning did not show a clear pattern, it seemed to have increased a bit more in the control than in the intervention group over time.

[INSERT TABLE 2 HERE.]

The Pearson product-moment bivariate correlations for the whole sample (n = 374) are presented in Table 3. Each variable was measured at five different points in time (t0–t4). Correlations ranged from -0.34 to 0.71. The items in all three SAL dimensions correlated positively with themselves over time, effect sizes being at least at the medium level ($r > |.30|$, see Cohen, 1988): r_{Deep} = 0.38–0.65, $r_{Surface}$ = 0.38–0.67, $r_{Organised}$ = 0.53–0.71.

[INSERT TABLE 3 HERE.]

Research question 1: Students' approaches to learning in two different instructional models

The analyses started with the selection of predictors for the multilevel growth models of the three DV's. Treatment (0 = control, 1 = intervention) was included as a predictor variable to the previously presented models for the deep, surface and organised (SAL) dimensions. A comparison of the models with LRT showed that treatment was a statistically significant predictor for the deep [$\chi^2(1)$ = 8.175, p = 0.004] and organised [$\chi^2(1)$ = 13.069, p < .001] models. However, an investigation of a possible interaction effect between time and treatment via LRT showed that such interaction was present only in the surface model [$\chi^2(1)$ = 3.987, p = 0.046]. This finding led to the inclusion of treatment as a predictor also in the surface model. Gender (0 = female, 1 = male) was also added as a predictor in all three SAL models, although it had a statistically significant association only with the surface approach to learning (this finding will be discussed later).

Deep approach to learning

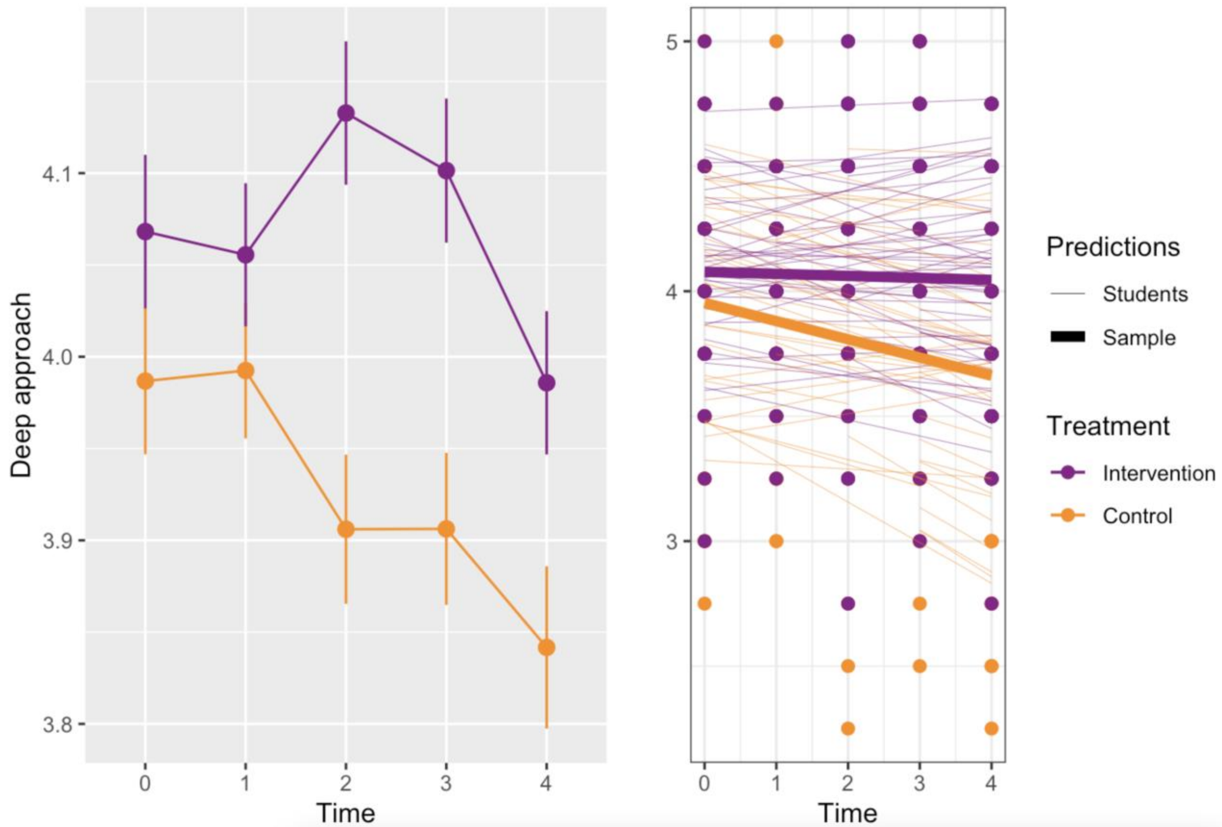
The results of the mixed-effect growth curve modeling of the deep approach to learning are presented in Table 4 (*sjPlot* package, see Lüdecke, 2016). Model 0 represents an unconditional means model while Model 1 has Time as a growth component with random intercepts and Model 2 has both random intercepts and slopes. The next two models are based on Model 2, but they have Treatment (Model 3) and Gender (Model 4) as predictors.

[INSERT TABLE 4 HERE.]

The ICC of Model 0 indicates that 55.3% of the variance in the deep approach to learning scores was attributed to student-to-student 'cross-sectional' variation (some were above, and some were below the average level), whereas 44.7% (1 - 0.553) of the variance was due to 'longitudinal' within-student measurement-to-measurement variation. Adding a linear time slope (in Model 2) increased between-student variation to 59.9%, indicating variations in the students' average values from the repeated measurements. Mean growth rate (time) was statistically significantly negative, indicating that deep approach to learning decreased over time. The variance of this growth rate (τ_{11} = 0.005, *C.I.* = 0.003 – 0.009) suggested the existence of individual differences in the rate of growth. Results showed no intercept-slope correlation (ρ_{01} = -0.05, C.I. = -0.34 – 0.25).

Treatment was also a statistically significant positive predictor (Model 3), highlighting the fact that participating in the intervention group (coded as '1') had a positive impact on the students' deep approach to learning (the control group was coded as '0'). Although the students' deep approach to learning declined over time, those in the intervention group remained on a higher level than the students in the control group (Figure 2). As Model 4 shows, there was no interaction effect between time and treatment and gender was not associated with deep approach to learning.

Figure 2. Development of deep approach to learning over time



Note: Left hand side of the figure displays the observed values of the intervention (n = 174) and control (n = 200) groups. Right-hand side contains predicted values (based on Model 3 in Table 4) of the random sample of 60 students.

Power and effect size were analysed with *simr::powerSim* (Green & MacLeod, 2016) using 200 simulations (α = .05). The mixed-effects model used in the power analyses had each SAL dimension in turn as a dependent variable, Treatment as a fixed predictor and Time as random predictor (both intercepts and slopes were allowed to vary for each student). The estimated power for the deep approach model was found to be 83.0% (*C.I.* = 77.1% – 87.9%), exceeding the desired level of 80.0% (see Cohen, 1988). However, the effect size (d = 0.12) was below the level of small effect (d = 0.20, see Cohen, 1988).

Surface approach to learning

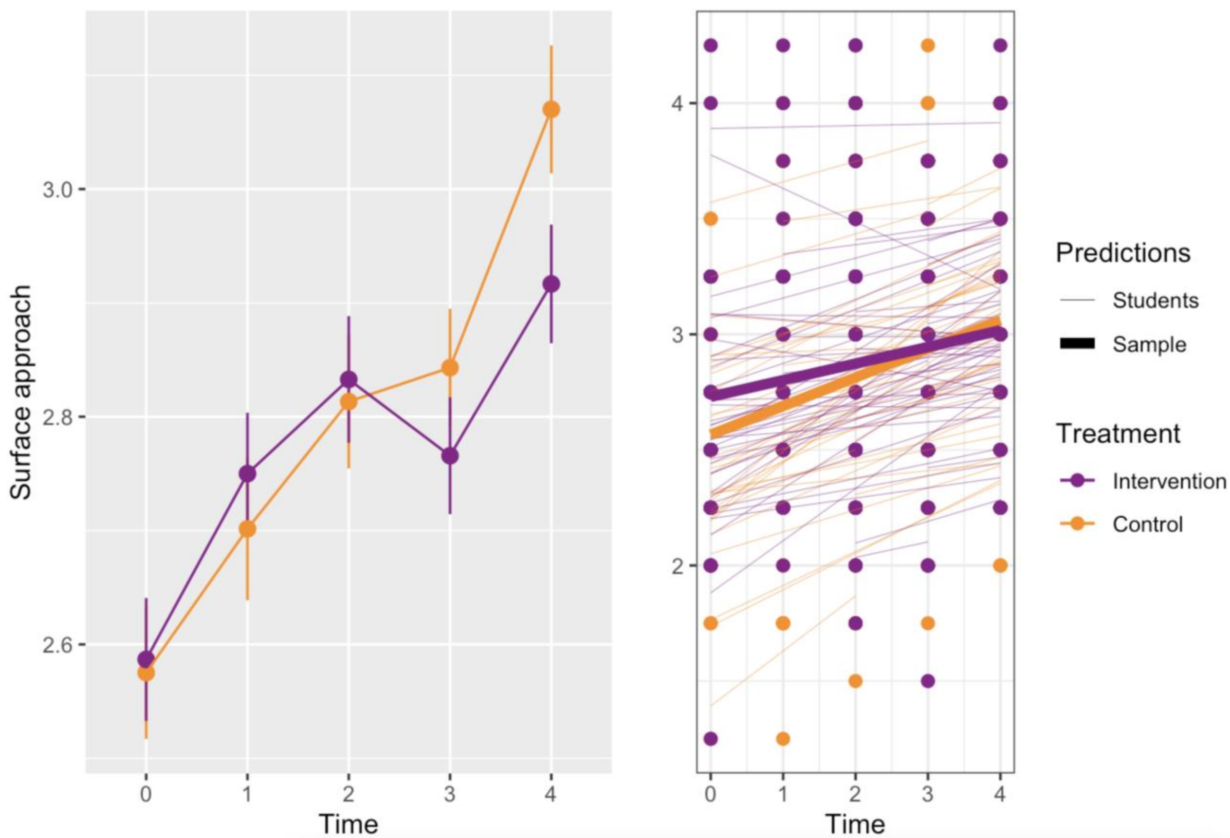
The results of the mixed-effect growth curve modeling of the surface approach to learning are presented in Table 5.

[INSERT TABLE 5 HERE.]

The mean growth rate (time) was statistically significantly positive, indicating that surface approach to learning increased over time (Model 2 in Table 5). The variance of this growth rate (τ_{11} = 0.015, *C.I.* = 0.010 – 0.021) was greater compared to that of the deep approach (τ_{11} = 0.005, see Model 2 in Table 4), suggesting that students differed more over time in their surface approach than in their deep approach. Moreover, the strong negative intercept-slope correlation (ρ_{01} = -0.41, C.I. = -0.56 – -0.23) in Model 2 underscored that the students with lower initial levels of surface approach to learning had steeper rates of increase over time than those with higher initial levels.

Treatment was included as a predictor in Model 3. It was a statistically significant negative predictor of surface approach only when interacting with time (Model 4). That is, the surface approach reached higher levels over time in the control group than in the intervention group (Figure 3). These findings should be interpreted with caution, as the estimated power of predicting the change in surface approach over time was low (11.0%, *C.I.* = 7.0% – 16.2%) alongside with a nonexistent effect size (d = - 0.04).

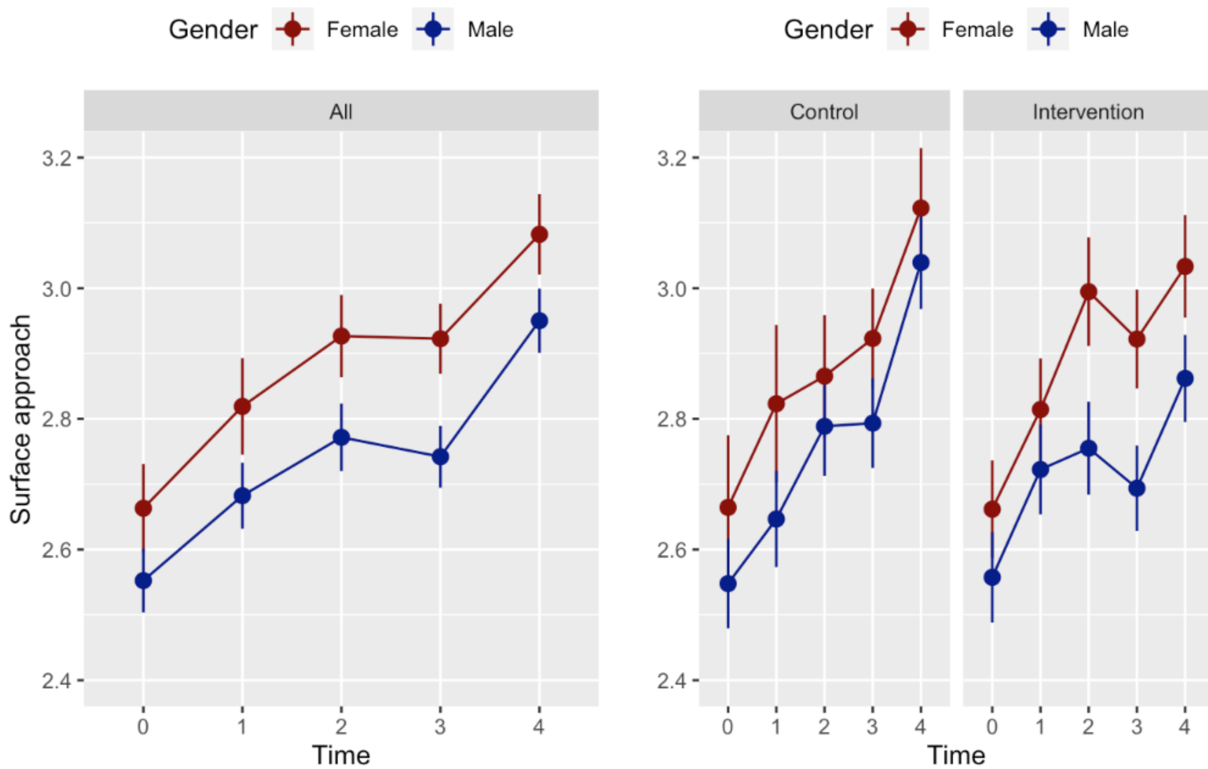
Figure 3. Development of surface approach to learning over time



Note: Left hand side of the figure displays the observed values of the intervention (n = 174) and control (n = 200) groups. Right-hand side contains predicted values (based on Model 3 in Table 5) of the random sample of 60 students for both groups.

As Model 4 in Table 5 shows, a gender effect was present. Negative estimate indicates that female students' (coded with value 0) surface approach stayed on a higher level than male students' (coded with value 1) surface approach over time (see Figure 4). The right-hand side of Figure 4 shows that this result was present in both control and intervention groups.

Figure 4. Development of surface approach to learning by gender over time



Note: Left hand side of the figure displays the observed values of the whole sample (*n* = 374) over time. The right-hand side contains the observed values for the intervention (*n* = 174) and control (*n* = 200) groups over time.

Organised studying

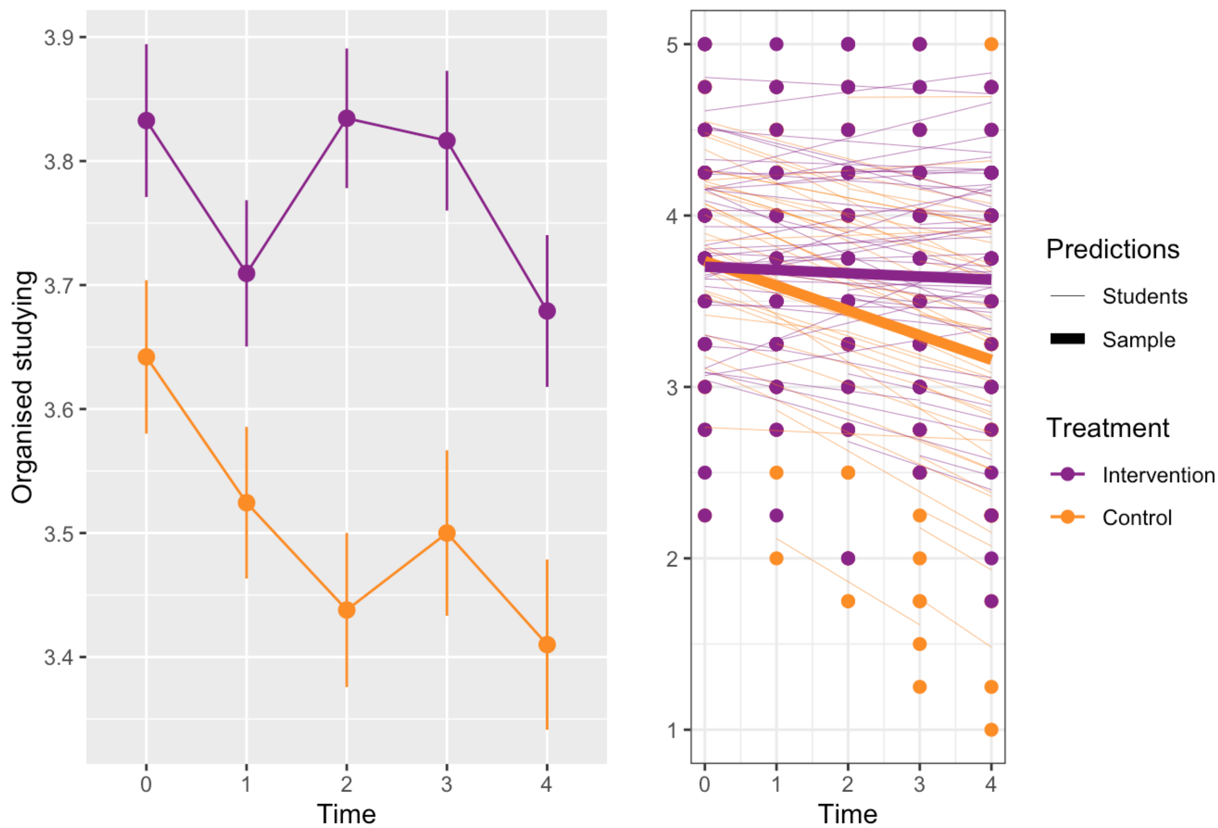
The results of the mixed-effect growth curve modeling of the organised studying are presented in Table 6. The variance components of Model 0 indicate that the within-student (unexplained) variation σ^2 (32.3% of the total variation $\sigma^2 + \tau_{00}$ = 0.589) was clearly smaller than those for deep approach (44.7%) and surface approach (41.0%). Contrary to the findings related to the deep and surface approaches, the ICC (0.677) for organised studying demonstrated that a greater portion (67.7%) of the variance was related to between-student variation, while a smaller portion of the variance (32.3%) was related to within-student variation. Therefore, the differences due to data clustering may be dominantly related to the disparities between the students' levels of organised studying (some are better organised than others) rather than differences in their repeated measurement profiles.

[INSERT TABLE 6 HERE.]

The mean growth rate (time) was statistically significantly negative, indicating that organised studying decreased over time (Model 1). Model 2 shows that the variance of this growth rate (τ_{11} = 0.012, *C.I.* = 0.007 – 0.022) was close to that of surface approach (τ_{11}= 0.015, see Model 2 in Table 5) but clearly larger compared to that of deep approach (τ_{11} = 0.005, see Model 2 in Table 4), suggesting that students differed more over time in their organised and surface approaches than in their deep approach.

The results related to Model 3 (Table 6) showed that treatment was a statistically significant positive predictor of organised studying. That is, organised studying remained quite stable over time in the intervention group, while it clearly declined in the control group (Figure 5). The estimated power for the organised studying model was high (94.0%, *C.I.* = 89.8% – 96.9%), but the effect size remained quite small (*d* = 0.25).

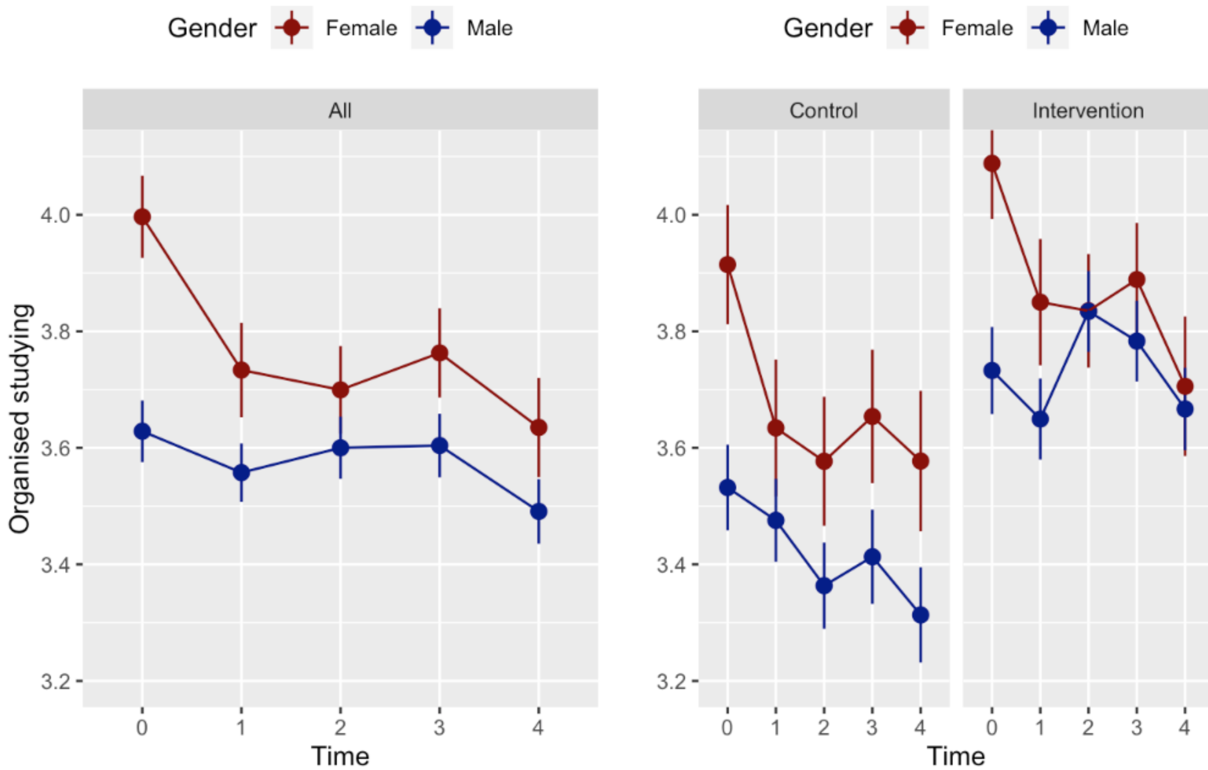
Figure 5. Development of organised studying over time



Note: The left-hand side of the figure displays the observed values of the intervention (*n* = 174) and control (*n* = 200) groups. The right-hand side contains predicted values (based on Model 3 in Table 6) of the random sample of 60 students for both groups.

Model 4 in Table 6 shows that there was no interaction between time and intervention. However, a negative estimate for gender indicated that female students' (coded with value 0) organised studying stayed on a higher level than male students' (coded with value 1) organised studying over time (see Figure 6). The right-hand side of the Figure 6 shows that this result was more clearly present in the control group than in the intervention group.

Figure 6. Development of organised studying by gender over time



Note: Left-hand side of the figure displays the observed values of the whole sample (n = 374) over time. The right-hand side contains the observed values for the intervention (n = 174) and control (n = 200) groups over time.

Research question 2: Students approaches to learning before and during the pandemic

The second research question was answered by analysing the data at each of the five timepoints using multilevel modeling. The first four measurements (t0–t3) took place before the pandemic, and the fifth (t4) was taken during the pandemic. The predictors of the DV variables (Deep, Surface, Organised) in all the models presented below were Treatment (0 = control, 1 = intervention) and Gender (0 = female, 1 = male). As the longitudinal data was split into five timepoints, the Time variable was not used in the analysis.

Deep approach to learning

Table 7 presents the multilevel modeling results of the deep approach to learning at five timepoints. The treatment effect stayed positive over time, indicating that in the intervention group, the level of deep approach increased over time more than in the control group. However, the effect was statistically significant from the third measurement onwards (t2, after the second course). The result may be due to a lower number of participants in the first two timepoints than in the last three timepoints. Results showed no gender effect.

[INSERT TABLE 7 HERE.]

To further elaborate on these findings, contrasts related to time were analysed with the *emmeans* package (Lenth, 2021), together with the *lme* ‘SAS.contr’ and ‘intervals’ functions (Pinheiro & Bates, 2000). The last (fifth) measurement (t4, during pandemic) was set as the ‘baseline’ level of deep approach, and the four previous scores (t0–t3) were compared against it. The results for the whole sample showed that all measurements (estimated marginal means) prior to the pandemic were higher (t0 = 4.00; t1 = 3.98; t2 = 3.99; t3 = 3.99) and deviated statistically significantly from the baseline (t4) value of 3.91. This declining trend was best explained with both linear and quadratic change over time. Results with the control group data showed that only the first two measurements (t0 = 3.96; t1 = 3.95) were statistically significantly higher than the baseline (t4) value of 3.83. This indicates that this linear decline of the deep approach scores had started well before the pandemic restrictions took place. However, the analyses of the intervention group data showed both quadratic and cubic change over time: the deep approach to learning was inclining until the beginning of the pandemic. Estimated marginal means of the last two measurements (prior pandemic) at timepoints t2 (4.10) and t3 (4.07) were statistically significantly higher than at t4 during the pandemic (3.98).

Surface approach to learning

Investigations of the multilevel modeling results of the surface approach to learning at the five timepoints (Table 8) showed no treatment effect after the interaction with time had been removed (cross-sectional data). However, it became evident that the surface approach was higher in the intervention group than in the control group at the first three timepoints. Thereafter, it reached a higher level in the control group than in the intervention group (Table 8 and Figure 3). Surface approach was higher for the female students at all timepoints, but statistically significant difference was found only in the fourth measurement (t3, after the third course).

[INSERT TABLE 8 HERE.]

The results of the contrast analyses with the whole data highlighted that the surface approach estimated marginal means (t0 = 2.65, t1 = 2.79, t2 = 2.88, t3 = 2.85) prior to the last measurement (t4 = 3.06) were statistically significantly lower than at the last measurement, following both linear and cubic inclining trends.

Organised studying

Table 9 presents the multilevel modeling results of the organised studying at the five timepoints. The treatment association was statistically significant (positive) at all timepoints, indicating that the intervention group systematically rated their organised studying higher than the control group The association became stronger after the second course (timepoints t2–t4). Organised studying was higher for the female students at all timepoints, but statistically significant difference was found only in the first measurement (t0, before the first course).

[INSERT TABLE 9 HERE.]

Positive associations with treatment were further investigated with contrast analyses where the four (prior pandemic) measurements (t0–t3) were compared to the fifth measurement t4 (during pandemic). An investigation of the polynomial time trends (linear, quadratic, cubic, quartic) for the full data showed that although organised studying developed dominantly over time in linear way, a cubic trend was also statistically significant. The control group manifested a declining (negative) mostly linear trend over time, where the first (3.63) and last (3.38) measurements deviated statistically significantly ($p < 0.001$). By contrast, the intervention group demonstrated a mixture of linear, cubic and quartic trends over time. After the pandemic restrictions, a drop was apparent at t4 (during pandemic) compared to the previous measurements at t0 (t4–t0 = 0.17, p = 0.004) and t2 (t4–t2 = 0.19, p = 0.001).

Discussion

We examined engineering students' approaches to learning through a sequence of four mathematics courses. The students were divided into two groups that both received student-centred instruction. For the control group, the student-centred elements were added to a traditional lecture-based framework. The intervention group's instructional model was based on flipped learning and was designed from the point of view of student-centred learning. During the fourth course, the COVID-19 forced all teaching to shift to online format. In this study, we analysed students' approaches to learning in the two instructional models, both before and during the pandemic.

Before and during the pandemic, students in the intervention group reported using more favourable approaches to learning than the control group. Although the levels of deep approach to learning and organised studying declined over time, the students in the intervention group remained on a higher level than those in the control group. The surface approach to learning increased over time, but the students in the intervention group remained on a lower level than the students in the control group. It seems that the impact of the treatment effect came with a delay, as the surface approach of the intervention group declined only by measurement point 3. It may be that it took time for the students to adjust to a new instructional model.

One explanation for the higher levels of deep approach learning in the intervention group is collaborative activities and peer support. In previous studies, these factors have been linked to high levels of deep approach to learning (e.g., Coertjens et al., 2016; Lahdenperä et al., 2018; Waters & Johnston, 2004). Also, students in the intervention group were given specific tasks that aimed for conceptual understanding. These tasks may have helped students understand ideas for themselves and examine the big picture, which are skills characteristic of the deep approach to learning (Entwistle & Peterson, 2004). Finally, the fact that students were expected to be active and take responsibility for their own learning in the intervention group might explain the high level of deep-approach learning (Wilson & Fowler, 2005).

The high level of organised studying in the intervention group may be related to how the instructional model supported study management. Students were required to make study schedules for themselves, and study skills were discussed during the prime-time sessions with teachers. The students in the intervention group self-assessed their competencies throughout the courses, which made them reflect on their own learning and may have fostered organised studying (Clark, 2012).

Gender differences were detected in students' approaches to learning. In general, female students had higher levels of surface approach to learning and organised studying than male students. It appeared that male students benefited more from the intervention than female students. Their level of surface approach to learning was low and their level of organised studying rose almost to the level of female students in the intervention group. However, when interpreting the gender differences between participants, it should be remembered most students were male (n = 247, 66.3%).

The teaching restrictions caused by the COVID19 pandemic forced the teachers of both instructional models to redesign all teaching to remote form practically overnight. It should be noted that since neither of the models was originally designed as online teaching, compromises had to be done. In both groups, exercise sessions were canceled, and the tutoring lab was taught online. The main difference in contact teaching between the two groups was that in the control group the lectures were canceled and replaced by online videos, whereas in the intervention group the prime-time sessions continued. This meant that the control group had less contact teaching, and their schedule was less structured compared to the intervention group. In both groups, assignments were submitted on the online learning platform and self- and peer-assessed, which was new for the control group, but not for the intervention group. All in all, fewer changes in the teaching arrangements were needed in the intervention group compared to the control group.

The COVID19 pandemic had a negative impact on students' approaches to learning. During the pandemic, students reported a lower level of deep approach to learning and organised studying than before the pandemic. One explanation for the drop in deep approach to learning are meeting restrictions and distance learning that halted access to the structures that fostered peer support, such as exercise groups, providing fewer opportunities for students to indulge in mathematical discussions (see e.g., Coertjens et al., 2016; Lahdenperä et al., 2018; Waters & Johnston, 2004). The lack of regular meetings also may explain the decline in organised studying. Even though the deep approach to learning decreased, the level of deep approach to learning remained higher in the intervention group. This might be due to the regular prime-time meetings that continued in the intervention group in which the students met with a teacher and a fixed small group of peers.

The surface approach to learning increased in both student groups during the pandemic. When teaching practices changed due to the pandemic, it happened unexpectedly. Both teachers and students had to learn how to use new digital tools in a couple of days. Also, teachers had only a few days to plan new teaching practices. In addition, the overall uncertainty that the pandemic caused, as well as meeting restrictions, generated stress and anxiety among students. Big workload and stress have been linked to increased levels of surface approach to learning in the literature (Cheung et al., 2020).

Limitations of the study

The control and intervention groups were not chosen randomly but based on students’ study programmes. It is known that approaches to learning are related to individual students’ characteristics (Baeten et al., 2010). To take this into account, the first measurement was done at the very beginning of the first course. No significant difference in the students’ approaches to learning between the control and intervention groups was observed in the first measurement. However, the cultures of different engineering study programmes (Lattuca, 2010) may have influenced students’ approaches to learning after the initial measurement.

Altogether, six teachers led the four courses examined in this manuscript. Teachers' approaches to teaching may influence students' approaches to learning (Baeten et al., 2010). However, all the responsible teachers were experienced and had pedagogical qualifications.

The data were self-reported, but the questionnaire has been validated in the Finnish, UK and Denmark higher education contexts (Parpala, 2010; Herrman et al., 2017). One limitation of the present study is the low internal consistency value (α = .52) of the deep approach to learning dimension. This may be because the present study had a significantly smaller number of respondents (N = 374) than the three previous studies (Finland N = 2509, UK N = 2710, Denmark N = 4377). The participants comprised a selected group in the sense that they were among those who did not drop out from the courses and answered all five questionnaires.

Implications for teaching and future research

Our results suggest that the instructional model used for the intervention group supported favourable approaches to learning better than the one used for the control group, both before and during the pandemic. Based on prior research, we hypothesise that important elements that supported deep approach to learning in the intervention group include collaborative work, fixed small groups in which the students worked, conceptual tasks and self-assessment. All in all, the results indicate that, when implementing student-centred teaching, it is more efficient to design new structures than just add student-centred elements to traditional structures such as lectures (see Lahdenperä et al., 2019).

In the future, more information is needed on which elements of the instructional model exerted the positive outcome in the intervention group. Collecting qualitative data, e.g., in the form of interviews and observations, could provide more information on students' approaches to learning and the characteristics of instructional models that foster favourable approaches to learning.

Disclosure statement

The authors report there are no competing interests to declare.

References

Asikainen, H., Parpala, A., Lindblom-Ylänne, S., Vanthournout, G., & Coertjens, L. (2014). The development of approaches to learning and perceptions of the teaching-learning environment during bachelor level studies and their relation to study success. *Higher Education Studies, 4*(4), 24–36. https://doi.org/10.5539/hes.v4n4p24

Baeten, M., Kyndt, E., Struyven, K., & Dochy, F. (2010). Using student-centred learning environments to stimulate deep approaches to learning: Factors encouraging or discouraging their effectiveness. *Educational Research Review*, *5*, 243–260. https://doi.org/10.1016/j.edurev.2010.06.001

Baeten, M., Struyven, K., & Dochy, F. (2013). Student-centred teaching methods: Can they optimise students' approaches to learning in professional higher education?. *Studies in Educational Evaluation*, *39*(1), 14–22. https://doi.org/10.1016/j.stueduc.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.48550/arXiv.1406.5823

Baytiyeh, H., & Naja, M. K. (2017). Students' perceptions of the flipped classroom model in an engineering course: a case study. *European Journal of Engineering Education*, *42*(6), 1048–1061. https://doi.org/10.1080/03043797.2016.1252905

Biggs, J. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *British journal of educational psychology*, 63(1), 3–19. https://doi.org/10.1111/j.2044-8279.1993.tb01038.x

Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach.* Springer.

Cheung, K., Yip, T. L., Wan, C. L. J., Tsang, H., Zhang, L. W., & Parpala, A. (2020). Differences in study work-load stress and its associated factors between transfer students and freshmen entrants in an Asian higher education context. *PLoS ONE*, *15*(5). https://doi.org/10.1371/journal.pone.0233022

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, *24*, 205–249. https://doi.org/10.1007/s10648-011-9191-6

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Second edition. Erlbaum.

Coertjens, L., Vanthournout, G., Lindblom-Ylänne, S., & Postareff, L. (2016). Understanding Individual Differences in Approaches to Learning Across Courses: A Mixed Method Approach. *Learning and Individual Differences*, *51*, 69–80. https://doi.org/10.1016/j.lindif.2016.07.003

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction*, *13*(5), 533–568. https://doi.org/10.1080/01443410303225

Dolmans, D. H., Loyens, S. M., Marcq, H., & Gijbels, D. (2016). Deep and surface learning in problem-based learning: a review of the literature. *Advances in health sciences education*, 21(5), 1087–1112. https://doi.org/10.1007/s10459-015-9645-6

Du, X., U. Ebead, S. Sabah, J. Ma, and K. K. Naji. 2019. "Engineering Students' Approaches to Learning and Views on Collaboration: How do Both Evolve in a PBL Environment and What Are Their Contributing and Constraining Factors?" *EURASIA Journal of Mathematics, Science and Technology Education*, *15*(11), em1774. https://doi.org/10.29333/ejmste/106197

Entwistle, N., McCune, V., & Hounsell, J. (2003). Investigating ways of enhancing university teaching-learning environments: Measuring students' approaches to studying and perceptions of teaching. In E. De Corte, L. Verschaffel, N. Entwistle, & J. van Merrienboer (Eds.), *Powerful learning environments: Unravelling basic components and dimensions* (pp. 89–108). Elsevier Science Limited.

Entwistle, N. & McCune, V. (2004). The conceptual bases of study strategy inventories. *Educational Psychology Review, 16*(4), 325–345.

Entwistle, N., McCune, V., & Scheja, M. (2006). Student learning in context: understanding the phenomenon and the person. *Instructional psychology: Past, present and future trends*, 131-148.

Entwistle, N. J., & Peterson, E. R. (2004). Conceptions of learning and knowledge in higher education: Relationships with study behaviour and influences of learning environments. *International journal of educational research*, *41*(6), 407–428. https://doi.org/10.1016/j.ijer.2005.08.009

Entwistle, N., & Ramsden, P. (1983). *Understanding student learning*. Croom Helm.

Evans, S., Alkan, E., Bhangoo, J. K., Tenenbaum, H., & Ng-Knight, T. (2021). Effects of the COVID-19 lockdown on mental health, wellbeing, sleep, and alcohol use in a UK student sample. *Psychiatry Research*, *298*, 113819. https://doi.org/10.1016/j.psychres.2021.113819

Finch, W. H., Bolin, J. E., & Kelley, K. (2014). *Multilevel Modeling Using R*. Chapman and Hall/CRC.

Fox, J., & Weisberg, S. (2019). *An {R} Companion to Applied Regression*, Third Edition. Thousand Oaks CA: Sage.

Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Green, P., & MacLeod, C. J. (2016). simr: an R package for power analysis of generalised linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498.

Haarala-Muhonen, A., Ruohoniemi, M., Parpala, A., Komulainen, E., & Lindblom-Ylänne, S. (2017). How do the Different Study Profiles of First-year Students Predict their Study Success, Study Progress and the Completion of Degrees? *Higher Education*, *74*(6), 949–962. https://doi.org/10.1007/s10734-016-0087-8

Harrison, E., Drake, T., & Ots, R. (2021). *finalfit: Quickly Create Elegant Regression Results Tables and Plots when Modelling*. R package version 1.0.3. https://CRAN.R-project.org/package=finalfit

Herrmann, K. J., Bager-Elsborg, A., & Parpala, A. (2017). Measuring perceptions of the learning environment and approaches to learning: validation of the Learn questionnaire. *Scandinavian*

Journal of Educational Research, *61*(5), 526–539. https://doi.org/10.1080/00313831.2016.1172497

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, *42*(8), 1–28. https://doi.org/10.18637/jss.v042.i08

Hoffman, L. (2015). *Longitudinal analysis: Modeling within-person fluctuation and change.* Routledge.

Hox, J. (2010). *Multilevel Analysis Techniques and Applications*, Second Edition. Routledge.

Karabulut-Ilgu, A., Jaramillo Cherrez, N., & Jahren, C. T. (2018). A systematic review of research on the flipped learning method in engineering education. *British Journal of Educational Technology*, *49*(3), 398–411. https://doi.org/10.1111/bjet.12548

Koskinen, P., Lämsä, J., Maunuksela, J., Hämäläinen, R., & Viiri, J. (2018). Primetime learning: collaborative and technology-enhanced studying with genuine teacher presence. *International journal of STEM education*, *5*(1), 20. https://doi.org/10.1186/s40594-018-0113-8

Lahdenperä, J., Postareff, L., & Rämö, J. (2019). Supporting quality of learning in university mathematics: A comparison of two instructional designs. *International Journal of Research in Undergraduate Mathematics Education*, *5*(1), 75–96. https://doi.org/10.1007/s40753-018-0080-y

Lattuca, L. R., Terenzini, P. T., Harper, B. J., & Yin, A. C. (2010). Academic environments in detail: Holland's theory at the subdiscipline level. *Research in Higher Education*, *51*(1), 21–39. https://doi.org/10.1007/s11162-009-9144-9

Laursen, S. L., Hassi, M. L., Kogan, M., & Weston, T. J. (2014). Benefits for women and men of inquiry-based learning in college mathematics: A multi-institution study. *Journal for Research in Mathematics Education*, *45*(4), 406–418. https://doi.org/10.5951/jresematheduc.45.4.0406

Lenth, R. (2021). emmeans: *Estimated Marginal Means, aka Least-Squares Means.* R package version 1.8.4-1. https://CRAN.R-project.org/package=emmeans

Leung, M. Y., Lu, X., Chen, D., & Lu, M. (2008). Impacts of teaching approaches on learning approaches of construction engineering students: A comparative study between Hong Kong and Mainland China. *Journal of Engineering Education*, *97*(2), 135–145. https://doi.org/10.1002/j.2168-9830.2008.tb00963.x

Lindblom-Ylänne, S., Parpala, A. M., & Postareff, L. (2019). What Constitutes the Surface Approach to Learning in the Light of New Empirical Evidence? *Studies in Higher Education*, *44(12)*, 2183–2195. https://doi.org/10.1080/03075079.2018.1482267

Lüdecke, D. (2016). *SjPlot: Data visualization for statistics in social science*. http://CRAN.R-project.org/package=sjPlot.

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I – Outcome and process. *British Journal of Educational Psychology*, *46*(1), 4–11. https://doi.org/10.1111/j.2044-8279.1976.tb02980.x

Minbashian, A., Huon, G. F., & Bird, K. D. (2004). Approaches to studying and academic performance in short-essay exams. *Higher Education*, *47*(2), 161–176. https://doi.org/10.1023/B:HIGH.0000016443.43594.d1

Mushquash, A. R., & Grassia, E. (2020). Coping during COVID-19: examining student stress and depressive symptoms. *Journal of American college health*, 1–4. https://doi.org/10.1080/07448481.2020.1865379

Parpala, A. (2010). *Exploring the experiences and conceptions of good teaching in higher education: Development of a questionnaire for assessing students‘ approaches to learning and experiences of the teaching-learning environment* (Doctoral dissertation). Studies in Educational Sciences 230. Helsinki University Press.

Parpala, A., & Lindblom-Ylänne, S. (2012). Using a research instrument for developing quality at the university. *Quality in Higher Education*, *18*(3), 313–328.

Pinheiro, J., & Bates, D. (2000). *Mixed Effects Models in S and S-Plus*. Springer.

Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2021). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-152. https://CRAN.R-project.org/package=nlme.

Polanco, R., Calderón, P., & Delgado, F. (2004). Effects of a problem-based learning program on engineering students' academic achievements in a Mexican university. *Innovations in Education and Teaching International*, *41*(2), 145–155. https://doi.org/10.1080/1470329042000208675

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, *93*(3), 223–231. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Raudenbush, S., & Bryk, A. (1986). A hierarchical model for studying school effects. *Sociology of Education*, *59*, 1–17. https://doi.org/10.2307/2112482

RStudio Team (2016). *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA. http://www.rstudio.com.

Rytkönen, H., Parpala, A., Lindblom-Ylänne, S., Virtanen, V., & Postareff, L. (2012). Factors affecting bioscience students' academic achievement. *Instructional Science*, *40*(2), 241–256. https://doi.org/10.1007/s11251-011-9176-3

Struyven, K., Dochy, F., Janssens, S., & Gielen, S. (2006). On the dynamics of students' approaches to learning: The effects of the teaching/learning environment. *Learning and instruction*, *16*(4), 279–294. https://doi.org/10.1016/j.learninstruc.2006.07.001

Swayne, D. F., & Buja, A. (1998). *Missing Data in Interactive High-Dimensional Data Visualization*. https://www.researchgate.net/publication/2758672_Missing_Data_in_Interactive_High-Dimensional_Data_Visualization.

Talbert, R. (2017). *Flipped learning: A guide for higher education faculty.* Stylus Publishing.

Tierney, N., Cook, D., McBain, M,, & Fay, C. (2021). *naniar: Data Structures, Summaries, and Visualisations for Missing Data*. R package version 0.6.1. https://CRAN.R-project.org/package=naniar

van Buuren, S. (2018). *Flexible Imputation of Missing Data*, Second Edition. Chapman and Hall/CRC.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations. *R. Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Waters, L., & Johnston, C. (2004). Web-delivered, problem-based learning in organisational behaviour: a new form of CAOS. *Higher Education Research & Development*, *23*(4), 413–431. https://doi.org/10.1080/0729436042000276440

Wiley, J. F. (2020). multilevelTools: *Multilevel and Mixed Effects Model Diagnostics and Effect Sizes*. R package version 0.1.1. https://CRAN.R-project.org/package=multilevelTools

Wilson, K., & Fowler, J. (2005). Assessing the impact of learning environments on students’ approaches to learning: comparing conventional and action learning designs. *Assessment & Evaluation in Higher Education, 30*(1), 87–101. https://doi.org/10.1080/0260293042003251770

Yadav, A., Subedi, D., Lundeberg, M. A., & Bunting, C. F. (2011). Problem-based learning: Influence on students' learning in an electrical engineering course. *Journal of Engineering Education*, *100*(2), 253–280. https://doi.org/10.1002/j.2168-9830.2011.tb00013.x

Table 1. Teaching arrangements for two student groups, control and intervention, that were taught using two different student-centred instructional models.

	Before the pandemic	
	Control	**Intervention**
Contact teaching	lectures, exercise groups	prime-time meetings, exercise groups
Support	tutoring lab, basic skills support sessions	tutoring lab, basic skills support sessions
Materials	course materials, videos, computerised tasks, pen-and-paper tasks	course materials, videos, computerised tasks, pen-and-paper tasks, conceptual tasks
Collaboration	working with peers in the exercise groups and lectures	working with a fixed small group in prime-time meetings and exercise groups; completing group assignments
Assessment	weekly tasks, exam	weekly tasks, exam, peer assessment, self-assessment of tasks and learning objectives
	During the pandemic	
	Control	**Intervention**

Contact teaching		Prime-time meetings
Support	tutoring lab, basic skills support sessions	tutoring lab, basic skills support sessions
Materials	course materials, videos, computerised tasks, pen-and-paper tasks	course materials, videos, computerised tasks, pen-and-paper tasks
Collaboration		working with a fixed small group in prime-time meetings, completing group assignments
Assessment	weekly tasks, exam, self-assessment and peer assessment of tasks	weekly tasks, exam, peer assessment, self-assessment of tasks and learning objectives

Table 2. Descriptive statistics

	Deep (α = 0.52) *M*(*SD*)	Surface (α = 0.67) *M*(*SD*)	Organised (α = 0.79) *M*(*SD*)
Control (*n* = 200)			
t0	3.99(0.465)	2.58(0.671)	3.65(0.710)
t1	4.00(0.430)	2.70(0.726)	3.53(0.704)
t2	3.91(0.499)	2.81(0.720)	3.44(0.759)
t3	3.91(0.496)	2.84(0.622)	3.49(0.802)
t4	3.84(0.541)	3.07(0.689)	3.41(0.841)
Intervention (*n* = 174)			
t0	4.07(0.460)	2.59(0.594)	3.83(0.678)
t1	4.06(0.422)	2.75(0.579)	3.71(0.638)
t2	4.13(0.471)	2.83(0.671)	3.83(0.678)
t3	4.10(0.469)	2.77(0.614)	3.82(0.674)
t4	3.99(0.464)	2.92(0.619)	3.68(0.727)

Table 3. Bivariate correlations among deep and surface approaches to learning and organised studying

	Deep.t0	Deep.t1	Deep.t2	Deep.t3	Deep.t4	Surf.t0	Surf.t1	Surf.t2	Surf.t3	Surf.t4	Orga.t0	Orga.t1	Orga.t2	Orga.t3	Orga.t4
Deep.t0	—														
Deep.t1	0.48	—													
Deep.t2	0.49	0.47	—												
Deep.t3	0.45	0.47	0.65	—											
Deep.t4	0.38	0.47	0.58	0.57	—										
Surf.t0	-0.34	-0.14	-0.14	-0.17	-0.09	—									
Surf.t1	-0.06	-0.15	-0.11	-0.18	-0.11	0.61	—								
Surf.t2	-0.11	-0.11	-0.19	-0.25	-0.18	0.56	0.62	—							
Surf.t3	-0.18	-0.17	-0.30	-0.29	-0.23	0.51	0.56	0.67	—						
Surf.t4	-0.11	-0.10	-0.19	-0.21	-0.28	0.38	0.50	0.59	0.67	—					
Orga.t0	0.33	0.12	0.25	0.30	0.20	-0.27	-0.11	-0.23	-0.17	-0.17	—				
Orga.t1	0.17	0.12	0.20	0.17	0.26	-0.24	-0.22	-0.23	-0.19	-0.20	0.59	—			
Orga.t2	0.16	0.16	0.26	0.31	0.25	-0.13	-0.09	-0.22	-0.12	-0.13	0.61	0.71	—		
Orga.t3	0.27	0.13	0.33	0.37	0.30	-0.16	-0.10	-0.23	-0.18	-0.14	0.56	0.63	0.71	—	
Orga.t4	0.13	0.10	0.19	0.20	0.31	-0.04	-0.04	-0.24	-0.19	-0.27	0.53	0.63	0.67	0.67	—

Note: Deep = Deep approach to learning; Surf = Surface approach to learning; Orga = Organised studying.

Table 4. Changes over time and the treatment effect on deep approach to learning

	Model 0 *Deep ~ 1*			Model 1 *Deep ~ Time ri*			Model 2 *Deep ~ Time ris*			Model 3 *Deep ~ Time + Treatment ris*			Model 4 *Deep ~ Time * Treatment + Gender ris*		
Predictors	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*
(Intercept)	3.968 (3.926 – 4.009)	0.021	**<0.001**	4.009 (3.959 – 4.059)	0.025	**<0.001**	4.017 (3.970 – 4.065)	0.024	**<0.001**	3.961 (3.901 – 4.022)	0.031	**<0.001**	3.988 (3.902 – 4.075)	0.044	**<0.001**
Time				-0.019 (-0.032 – -0.006)	0.007	**0.004**	-0.023 (-0.037 – -0.008)	0.008	**0.003**	-0.022 (-0.037 – -0.008)	0.008	**0.003**	-0.034 (-0.054 – -0.013)	0.011	**0.001**
Treatment										0.119 (0.038 – 0.199)	0.041	**0.004**	0.080 (-0.015 – 0.175)	0.048	0.098
Gender													-0.012 (-0.098 – 0.074)	0.044	0.789
Time * Treatment													0.023 (-0.006 – 0.053)	0.015	0.121
Random Effects															
σ^2	0.107			0.107			0.095			0.095			0.096		
τ_{00}	0.133 sid			0.132 sid			0.115 sid			0.114 sid			0.113 sid		
τ_{11}							0.005 sid.Time			0.005 sid.Time			0.005 sid.Time		
ρ_{01}							-0.05 sid			-0.08 sid			-0.07 sid		
ICC	0.553			0.553			0.599			0.591			0.590		
N	373 sid			373 sid			373 sid			373 sid			373 sid		
Observations	1375			1375			1375			1375			1375		
Deviance	1452.628			1444.091			1420.978			1412.803			1410.342		
AIC	1458.628			1452.091			1432.978			1426.803			1428.342		

Note: ri = Random intercepts; *ris* = Random intercepts and slopes; *Est.* = Estimated values; *C.I.* = 95% Confidence Interval of the estimated values; *S.E.* = Standard error of estimated values; *p* = Statistical significance of the estimated values; *(Intercept)* = Overall DV average when predictors are set to zero; σ^2 = Variance of the level 1 residuals (within-students); τ_{00} = Variance of the cluster means (between-students); τ_{11} = Variance of the cluster (student) slopes over time; ρ_{01} = Overall correlation between students' intercepts and slopes; *ICC* = Adjusted intraclass correlation, proportion of the variance in DV explained by the random effects (clustering of the data); *N* = Number of students; *Observations* = Number of observations from students; *Deviance* = -2 · Log Likelihood, consistency of the model with the data; *AIC* = Akaike Information Criterion, predicting the power of predictors (Time, Treatment, Gender) on DV (Deep).

Table 5. Changes over time and impacts of treatment on surface approach to learning

	Model 0 *Surface ~ 1*			**Model 1** *Surface ~ Time ri*			**Model 2** *Surface ~ Time ris*			**Model 3** *Surface ~ Time + Treatment ris*			**Model 4** *Surface ~ Time * Treatment + Gender ris*		
Predictors	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*
(Intercept)	2.828 (2.769 – 2.887)	0.030	**<0.001**	2.641 (2.573 – 2.710)	0.035	**<0.001**	2.646 (2.573 – 2.719)	0.037	**<0.001**	2.650 (2.560 – 2.740)	0.046	**<0.001**	2.714 (2.586 – 2.842)	0.065	**<0.001**
Time				0.088 (0.072 – 0.105)	0.008	**<0.001**	0.086 (0.066 – 0.107)	0.010	**<0.001**	0.086 (0.065 – 0.106)	0.010	**<0.001**	0.107 (0.078 – 0.136)	0.015	**<0.001**
Treatment										-0.020 (-0.137 – 0.097)	0.060	0.734	0.073 (-0.071 – 0.217)	0.073	0.320
Gender													-0.163 (-0.287 – -0.040)	0.063	**0.010**
Time * Treatment													-0.041 (-0.081 – -0.001)	0.021	**0.045**
Random Effects															
σ^2	0.191			0.173			0.139			0.087			0.087		
τ_{00}	0.275 $_{sid}$			0.278 $_{sid}$			0.340 $_{sid}$			0.350 $_{sid}$			0.342 $_{sid}$		
τ_{11}							0.015 $_{sid.Time}$			0.018 $_{sid.Time}$			0.017 $_{sid.Time}$		
ρ_{01}							-0.41 $_{sid}$			-0.45 $_{sid}$			-0.44 $_{sid}$		
ICC	0.590			0.616			0.692			0.784			0.780		
N	373 $_{sid}$			373 $_{sid}$			373 $_{sid}$			373 $_{sid}$			373 $_{sid}$		
Observations	1375			1375			1375			1375			1375		
Deviance	2289.269			2186.263			2142.056			2122.352			2111.630		
AIC	2295.269			2194.263			2154.056			2154.352			2147.630		

Note: ri = Random intercepts; *ris* = Random intercepts and slopes; *Est.* = Estimated values; *C.I.* = 95% Confidence Interval of the estimated values; *S.E.* = Standard error of estimated values; *p* = Statistical significance of the estimated values; *(Intercept)* = Overall DV average when predictors are set to zero; σ2 = Variance of the level 1 residuals (within-students); τ00 = Variance of the cluster means (between-students); τ11 = Variance of the cluster (student) slopes over time; ρ01 = Overall correlation between students' intercepts and slopes; *ICC* = Adjusted intraclass correlation, proportion of the variance in DV explained by the random effects (clustering of the data); *N* = Number of students; *Observations* = Number of observations from students; *Deviance* = -2 · Log Likelihood, consistency of the model with the data; *AIC* = Akaike Information Criterion, predicting the power of predictors (Time, Treatment, Gender) on DV (Surface).

Table 6. Changes over time and impacts of treatment on organised studying

	Model 0 *Organised ~ 1*			Model 1 *Organised ~ Time ri*			Model 2 *Organised ~ Time ris*			Model 3 *Organised ~ Time + Treatment ris*			Model 4 *Organised ~ Time * Treatment + Gender ris*		
Predictors	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*
(Intercept)	3.565 (3.496 – 3.634)	0.035	**<0.001**	3.647 (3.569 – 3.725)	0.040	**<0.001**	3.654 (3.578 – 3.729)	0.038	**<0.001**	3.537 (3.440 – 3.634)	0.050	**<0.001**	3.666 (3.529 – 3.803)	0.070	**<0.001**
Time				-0.039 (-0.056 – -0.021)	0.009	**<0.001**	-0.043 (-0.063 – -0.022)	0.010	**<0.001**	-0.042 (-0.063 – -0.022)	0.010	**<0.001**	-0.054 (-0.082 – -0.025)	0.015	**<0.001**
Treatment										0.248 (0.115 – 0.382)	0.068	**<0.001**	0.220 (0.072 – 0.368)	0.075	**0.004**
Gender													-0.169 (-0.310 – -0.029)	0.072	**0.019**
Time × Treatment													0.022 (-0.019 – 0.063)	0.021	
Random Effects															
σ^2	0.190			0.160			0.160			0.160			0.161		
τ_{00}	0.399 $_{sid}$			0.381 $_{sid}$			0.362 $_{sid}$			0.350 $_{sid}$			0.340 $_{sid}$		
τ_{11}				0.014 $_{sid.Time}$			0.012 $_{sid.Time}$			0.012 $_{sid.Time}$			0.012 $_{sid.Time}$		
ρ_{01}				-0.15 $_{sid}$			-0.06 $_{sid}$			-0.09 $_{sid}$			-0.07 $_{sid}$		
ICC	0.677			0.704			0.724			0.715			0.711		
N	373 $_{sid}$			373 $_{sid}$			373 $_{sid}$			373 $_{sid}$			373 $_{sid}$		
Observations	1375			1375			1375			1375			1375		
Deviance	2401.214			2366.133			2350.083			2337.014			2330.434		
AIC	2407.214			2376.133			2362.083			2351.014			2348.434		

Note: ri = Random intercepts; *ris* = Random intercepts and slopes; *Est.* = Estimated values; *C.I.* = 95% Confidence Interval of the estimated values; *S.E.* = Standard error of estimated values; *p* = Statistical significance of the estimated values; *(Intercept)* = Overall DV average when predictors are set to zero; σ2 = Variance of the level 1 residuals (within-students); τ00 = Variance of the cluster means (between-students); τ11 = Variance of the cluster (student) slopes over time; ρ01 = Overall correlation between students' intercepts and slopes; *ICC* = Adjusted intraclass correlation, proportion of the variance in DV explained by the random effects (clustering of the data); *N* = Number of students; *Observations* = Number of observations from students; *Deviance* = -2 · Log Likelihood, consistency of the model with the data; *AIC* = Akaike Information Criterion, predicting the power of predictors (Time, Treatment, Gender) on DV (Organised).

Table 7. Multilevel model of deep approach to learning at five timepoints

	Before pandemic												During pandemic		
	Model t0			Model t1			Model t2			Model t3			Model t4		
Predictors	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*
(Intercept)	4.055 (3.937 – 4.173)	0.060	**<0.001**	4.046 (3.940 – 4.153)	0.055	**<0.001**	3.898 (3.789 – 4.006)	0.056	**<0.001**	3.932 (3.822 – 4.041)	0.056	**<0.001**	3.806 (3.694 – 3.918)	0.057	**<0.001**
Treatment	0.082 (-0.031 – 0.195)	0.058	0.156	0.064 (-0.041 – 0.169)	0.054	0.236	0.226 (0.116 – 0.337)	0.057	**<0.001**	0.197 (0.085 – 0.309)	0.057	**0.001**	0.141 (0.025 – 0.258)	0.059	**0.018**
Gender	-0.096 (-0.221 – 0.029)	0.064	0.134	-0.078 (-0.192 – 0.036)	0.058	0.183	0.013 (-0.104 – 0.130)	0.060	0.828	-0.040 (-0.158 – 0.078)	0.060	0.509	0.056 (-0.067 – 0.178)	0.062	0.372
N	253_{sid}			250_{sid}			294_{sid}			287_{sid}			291_{sid}		
Observations	253			250			294			287			291		
Deviance	319.838			276.915			405.216			394.954			425.672		
AIC	329.838			286.915			415.216			404.954			435.672		

Note: Est. = Estimated values; *C.I.* = 95% Confidence Interval of the estimated values; *S.E.* = Standard error of estimated values; *p* = Statistical significance of the estimated values; *N* = Number of students; *Observations* = Number of observations from students; *Deviance* = -2 · Log Likelihood, consistency of the model with the data; *AIC* = Akaike Information Criterion, predicting the power of predictors (Treatment, Gender) on DV (Deep).

Table 8. Multilevel model of surface approach to learning at five timepoints

	Before pandemic												During pandemic		
	Model t0			Model t1			Model t2			Model t3			Model t4		
Predictors	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*
(Intercept)	2.660 (2.497 – 2.824)	0.084	**<0.001**	2.796 (2.629 – 2.963)	0.085	**<0.001**	2.917 (2.761 – 3.073)	0.080	**<0.001**	2.953 (2.815 – 3.092)	0.071	**<0.001**	3.149 (3.004 – 3.294)	0.074	**<0.001**
Treatment	0.006 (-0.150 – 0.162)	0.080	0.939	0.050 (-0.114 – 0.214)	0.084	0.550	0.021 (-0.137 – 0.180)	0.081	0.794	-0.066 (-0.208 – 0.076)	0.073	0.362	-0.147 (-0.298 – 0.003)	0.077	0.056
Gender	-0.111 (-0.284 – 0.062)	0.088	0.211	-0.137 (-0.315 – 0.041)	0.091	0.133	-0.156 (-0.323 – 0.012)	0.086	0.071	-0.177 (-0.327 – -0.027)	0.077	**0.022**	-0.124 (-0.283 – 0.034)	0.081	0.125
N	253_{sid}			250_{sid}			294_{sid}			287_{sid}			291_{sid}		
Observations	253			250			294			287			291		
Deviance	484.138			499.492			616.665			531.552			575.724		
AIC	494.138			509.492			626.665			541.552			585.724		

Note: Est. = Estimated values; *C.I.* = 95% Confidence Interval of the estimated values; *S.E.* = Standard error of estimated values; *p* = Statistical significance of the estimated values; *N* = Number of students; *Observations* = Number of observations from students; *Deviance* = -2 · Log Likelihood, consistency of the model with the data; *AIC* = Akaike Information Criterion, predicting the power of predictors (Treatment, Gender) on DV (Surface).

Table 9. Multilevel model of organised studying at five timepoints

	Before pandemic												During pandemic		
	Model t0			**Model t1**			**Model t2**			**Model t3**			**Model t4**		
Predictors	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*	*Est. (C.I.)*	*S.E.*	*p*
(Intercept)	3.905 (3.731 – 4.080)	0.089	**<0.001**	3.648 (3.479 – 3.817)	0.086	**<0.001**	3.510 (3.348 – 3.672)	0.083	**<0.001**	3.612 (3.446 – 3.779)	0.085	**<0.001**	3.511 (3.337 – 3.684)	0.089	**<0.001**
Treatment	0.193 (0.027 – 0.360)	0.085	**0.024**	0.187 (0.020 – 0.353)	0.085	**0.029**	0.399 (0.235 – 0.564)	0.084	**<0.001**	0.325 (0.154 – 0.495)	0.087	**<0.001**	0.277 (0.096 – 0.457)	0.092	**0.003**
Gender	-0.370 (-0.554 – -0.185)	0.094	**<0.001**	-0.178 (-0.359 – 0.003)	0.092	0.055	-0.111 (-0.285 – 0.063)	0.089	0.214	-0.176 (-0.356 – 0.004)	0.092	0.057	-0.159 (-0.349 – 0.031)	0.097	0.103
N	253 $_{sid}$			250 $_{sid}$			294 $_{sid}$			287 $_{sid}$			291 $_{sid}$		
Observations	253			250			294			287			291		
Deviance	516.920			506.881			637.885			636.065			682.255		
AIC	526.920			516.881			647.885			646.065			692.255		

Note: Est. = Estimated values; *C.I.* = 95% Confidence Interval of the estimated values; *S.E.* = Standard error of estimated values; *p* = Statistical significance of the estimated values; *N* = Number of students; *Observations* = Number of observations from students; *Deviance* = -2 · Log Likelihood, consistency of the model with the data; *AIC* = Akaike Information Criterion, predicting the power of predictors (Treatment, Gender) on DV (Organised).