A moderate-structure cell biology course improves student performance but fails to alleviate destructive friction of a final comprehensive exam

Authors

IJsbrand Kramer^{*a}, Frédérique Pellerin^b and Xavier Nogues^c

^a) University of Bordeaux, UF Biologie, IECB – INSERM U1026, 2 rue Robert Escarpit, F-33607, Pessac Cedex, France

^b) BioStatistician Consultant, 96 Rue Camille Pelletan, 33400 Talence, France
(fs.pellerin@orange.fr)

^c) University of Bordeaux, UF Biologie, Avenue des Facultés, F-33405, Talence Cedex, France
(xavier.nogues@u-bordeaux.fr)

IJsbrand Kramer created the course formats and associated learning documents and taught the Cell Biology courses, Xavier Nogues and Frédérique Pellerin performed the statistical analyses.

*Corresponding author. ijsbrandkramer@gmail.com

Running title: moderate-structure course format and student achievement

Work telephone: 0033 (0)5 4000 8726

mobile: 0033 (0)6 2131 8220

Abstract

The transition from secondary to higher education remains problematic given low retention figures in science courses, in particular at open-enrolment universities. Adjustment to a new learning environment and dealing with the mass experience are factors at play. We looked for ways to ease the adjustment and to moderate the mass experience by creating a moderate-structure Cell Biology course, characterized by group-based activities, frequent in-course assessments and reduced weighting of the final exam score. Comparison of 4 years of low-structure with 4 years of moderate-structure courses, after correction for annual cohort ability, revealed that moderate-structure yields 8% higher grade points and 5% higher retention. However, the overall gain in performance was largely dependent on in-course scores and improvements were only weak for the final exam. The frequency of students underperforming on the final exam, relative to their in-course scores, increased enormously in moderate-structure courses, from 53 to 90.3%. We see this as a sign that for a substantial number of students, the final exam remained a destructive friction. As a result, we are still dealing with a population that has not assimilated a considerable portion of the cell biology knowledge and is starting the second year with significant gaps.

198 words

Key words; achievement, persistence, course structure, higher education, transition, cell biology, regulation, internalization, guidance

The Problematic of Transition from School to University

The transition from school to university is problematic as witnessed by low retention figures in science courses, in particular in “open enrolment” universities. This issue probably arose when higher education became commonplace in Western countries and first investigations into the causes of attrition dates back to the 1970s (Astin, 1972; Tinto, 1975). This and more recent studies revealed that three factors are at play: support from school, identity of learner and the way universities organize their first (introductory) year (Briggs, et al., 2002; Harvey et al., 2006; Hillman, 2005; Tinto, 1975). With respect to universities and in particular research-intensive universities with large enrolment courses, two things are special about the first year: 1) the adjustment to a new living and learning environment and, 2) dealing with the mass experience. With the latter is meant the experience of no longer being treated as an individual, of being confronted with a predominantly instructive, rather than facilitating, teaching approach (Ferrare, 2019) and the experience of being a (potential) problem for the teaching staff (because of the vast numbers of students and the competition with research time) (Brownell & Tanner, 2012; Harvey et al., 2006).

Structured Courses as a Means to Bridge the Gap Between School and University

Open enrolment universities in France have little to say about the identity of the student, nor about their number, but they have the possibility to facilitate adaptation and moderate the mass experience. Introducing “structured courses” (Eddy & Hogan, 2014) into the first year program is one of the possibilities to bring teaching in line with the students' school experience. These courses combine instructional learning with student-centered activities such as in-class quizzes and discussions with fellow students. The traditionally “high stake” final exam is replaced (in part) by frequent in-class assessments (Eddy & Hogan, 2014; Freeman et al., 2011; Gasiewski, et al., 2012; Haak et al., 2011). Indeed, numerous reports revealed that students appreciate structured courses (Sander et al., 2000) and that they had a positive impact on student engagement, resulting in better performances (Ebert-May et al., 1997; Eddy & Hogan, 2014; Freeman et al., 2007, 2011, 2014 (review); Haak et al., 2011; Knight & Wood, 2005; Tanner, 2013). It is well documented that the active-learning ingredient of structured courses, in particular peer discussions, activates the mind and benefits students’ (conceptual) understanding (Gasiewski, et al. 2012; Nelson, 2008; Smith, et al., 2009). It has also been made evident that students, and in particular underprepared students, hang on better in structured courses (more equity) (Haak, et al., 2011; Tanner 2013). One of the underlying explanations for enhanced

engagement is that the traditional instructive approach (often referred to as “gate-keeper” courses) frustrate students' inherent interests in science (Gasiewski, et al. 2012); it is not the subject but the way that it is presented that negatively affects motivation to learn.

A Moderate-Structure Cell Biology Course with Progressive Internalization of Regulation

With respect to retention, our first year introductory Cell Biology course made no exception. To change the situation, we have reconstructed the course into a *moderate-structure* course (Eddy & Hogan, 2014). In short, we subdivided the student population in groups of 6 students in order to foster social interaction with the aim of facilitating both peer instruction and the necessary social support (and as a consequence reduce the “mass experience”) (Cameron & Rideout, 2020; Yorke & Longden, 2008). Teacher instruction was interchanged with group-based exercises and quizzes with the use of electronic voting devices (“group-based clicker quizzes”) (Caldwell, 2007). In addition, we have reduced the weighting of the comprehensive final exam in favour of the scores obtained in clicker quizzes, in-course tests and a mid-course exam. Moreover, we created twelve on-line multimedia resources that covered the course content (composed of texts, images and animations). In addition, students had access to all slides and to a large bank of interactive multiple choice questions (some 200). In this way, we felt that students

Figure 1 Time-line of Course Activities: Lectures, Laboratory classes, In-class Clicker Quizzes, Tests and Exams

Note: This is a summary time-line of all courses over an 8-year period. Exactly what activities took place in low- and moderate-structure courses is described in Table 1.

had more control (autonomy, Ryan & Deci, 2000a, 2000b) over how they wanted to deal with the content of the course.

The moderate-structure course format, see figure 1 above, was also designed to orient teaching toward the learning process with particular attention to a progressive change of teacher control of learning. It starts with in-course activities, characteristic of strong teacher control, then moves to in-course tests, characteristic of shared teacher control, and the course terminates with a comprehensive exam with only loose teacher control. From the student perspective, the course transits from external guidance (teacher-based regulation), to shared guidance and terminates in internal guidance (students' self-regulation) (Ten Cate et al, 2004; Zimmerman, 2013). Importantly, this cognitive transition, from external to internal guidance, also pertains to the affective component and metacognitive component of learning. In this way, students should internalize all aspects of learning in a progressive way (Ten Cate, et al. 2004). This model of cognitive transition is based on the assumption that the quality of the regulation of learning is not a fixed trait, but instead is context-dependent. It assumes that learning environments play a role in how students regulate their learning (see also Kramer & Kusurkar, 2017). Driving students toward internal guidance, also described as self-regulation (Zimmerman, 2013), is important because it is positively associated with academic achievement (reviewed in Panadero, 2017) and it is also one the characteristics expected from university students when they enter the job market (Ten Cate et al., 2004).

We took the point of view that progressive internalization and social interaction should facilitate the necessary adaptation to the new, academic, learning environment and, as a consequence, improve student performance. We base this assumption on the hypothesis that progressive internalization should warrant an optimal level of congruence between learning and teaching (Vermunt & Verloop, 1999). Since we are dealing with a non-selective group of students, their *a priori* levels of regulation of learning must vary, with low-, medium-, and high-regulated students (Vermunt & Verloop, 1999). The in-class activities are considered particular beneficial for low-regulated students whereas the in-course tests are particularly suitable for intermediate-regulated students and, in addition, offer constructive friction for low-regulated learners. The final comprehensive exam should offer constructive friction for intermediate-regulated students and congruence for high-regulated learners (shown in table 4, Vermunt & Verloop, 1999). An important point to consider is that whatever course format, all teaching

approaches will, depending on the level of regulation of the learner, give rise to some destructive friction. In particular of concern is the final exam, which requires high internal guidance, and which consequently carries an element of destructive friction for low-regulated learners.

Destructive friction is also predicted for in-course activities with intermediate- and high-regulated students (Vermunt & Verloop, 1999). This, however, is less damaging as numerous studies reveal that students with higher levels of regulation can adapt better to unfavorable learning conditions than low-regulated learners (Beishuizen & Stoutjesdijk, 1999; Fryer & Vermunt, 2018).

The progressive internalization of the moderate-structure course, combined with group-based activities, could increase student engagement in a number of ways: 1) by extending time-on-task, 2) by distributing cognitive load, 3) by increasing the value of the learning task and 4) by enhanced psychological need satisfaction (including the necessary social integration). Through in-class activities, students are encouraged to learn in-class, something that some students would not normally do as passive bystanders in lectures. This may lead to an increase in the number of hours of engagement; the sum of time spent actively engaged in learning in- and outside the classroom. The incremental learning and assessment approach significantly reduces cognitive load (Merrienboer & Sweller, 2005) and this should be especially beneficial for students with low perceived self-efficacy (Bandura, 1977) or, as already mentioned, for students with low regulation (Vermunt & Verloop, 1999). The more student-centered teaching approach, in particular the social interaction, is likely to be translated in a higher perceived value of the course content (Townsend & Hicks, 1997). Both the access to comprehensive learning documents and the social interaction that follows from group activities should enhance the psychological need satisfaction necessary for full (self-determined) participation in learning tasks. Here we refer to the three psychological needs that, according to Self-Determination Theory, must be satisfied for volitional participation in a learning activity: a sense of autonomy (sense of agency), competence, and relatedness to peers (Ryan & Deci, 2000a, 2000b). Although moderated by personality traits and educational curricula (different disciplines), the degree to which these psychological needs are satisfied will also determine the degree to which students internalize their regulation. Internalization of motivation, being “self-determined”, is positively associated with student achievement and well-being (Chapter 8, Organismic Integration Theory, Ryan & Deci, 2017).

Research Questions.

Although this is to be expected, based on the numerous studies cited above, our primary concern was to determine whether student outcomes in courses with *moderate structure* are significantly better than in courses with *low structure*. The null hypothesis (H0) is that outcomes do not differ between *low-* and *moderate-structure* courses. The alternative hypothesis (H1) is that increased structure does affect study outcomes significantly and positively. In the case of significant improvement, we then wanted to know how students' results compared between the in-course assessments and the comprehensive final exam. As mentioned, the final exam in particular could be a destructive friction for some students. Do all students perform better in all knowledge assessments? Our null hypothesis (H0) is that all students move toward higher levels of internalization of learning and we find a proportional increase in all knowledge test scores (same outcome profile). The alternative hypothesis (H1) is that we see a disproportionate increase in the different assessment scores (different outcome profile). Such a shift in the profile would indicate at what level of the course students experience destructive friction.

We compared student performance in *low-* and *moderate-structure* courses over two periods of four years and involving a total of 3313 first year University students (Life, Earth & Environmental Sciences). The disadvantage of comparing different years and teaching methods is that you have more uncontrolled variables but by augmenting the number of years the likelihood that differences are the sole consequence of different teacher and/or different student populations is greatly diminished. With respect to course performance, we partially remedy this problem by correcting study performance for academic ability across cohorts. However, numerous uncontrolled variables is typically the problem of open- and high-enrolment education and, as teachers, we think it is important that the test ground has a high reality value (see also Andrews et al., 2011).

Methods

Participants Characteristics, Ethics and Assignment

The participants were first year university students with an average age of 19.7, SD=1.16. The population comprised 65.3 % women. The students enrolled in Life, Earth and Environmental Sciences (“Science de la Vie, de la Terre et de l’Environnement” or “SVTE”) at a French open-enrolment University. Enrolment requires a “baccalauréat”, a diploma equivalent of an American High School diploma or English GCE A-level, which is obtained with a minimal

grade-point average of 100/200. Each year, roughly 76% of the students came directly from secondary school (first entry) and 24% entered the first year at our university from another higher education institute or were repeaters (second entry). The students participated in a compulsory introductory Cell Biology course, held in the second semester (spring) of the 1st year. The course enrollment varied from 353 (in 2005) to 493 students (in 2016). Variation in student number is the consequence of, in order of importance, local demographic factors, students' first choices (of study) and the proportion of local candidates being accepted for selective higher education institutions (Grandes Écoles) in the region or elsewhere in France.

Our institution does not have an Institutional Review Board that would determine what teaching methods are or are not accepted. There are four conditions to be respected: face-to-face learning activities for all the allocated slots for lecture and laboratory-classes, an equivalent course content and equivalent teaching methods for all students enrolled in the same course (equality of chances of success), one mid-course exam (imposed from 2013 onwards) and an anonymous comprehensive final exam (for all years concerned).

Because the university requires identical course offerings with the same knowledge assessment, we could not separate the yearly cohorts into two populations, a “control” and “experimental”, and had to study the effect of the course structure in separate years. The years we chose for the study were solely determined by the completeness and comparability of the course offerings. There is a gap between 2006 and 2013, a period in which we developed multimedia resources and interactive self-evaluation quizzes. The spread of “control” and “experimental” classes over the years has the danger that we are not studying similar student populations. As mentioned earlier, we addressed this problem in two ways: first, by looking over a 4-year period and second, by correcting the cell biology results for semester grade-point average of each student.

Description of the Cell Biology Courses (Low- and Moderate-structure)

The current study concerns an introductory Cell Biology course which is offered once a year in the spring semester, January to end of April, to University students who enter Life, Earth and Environmental Sciences (Science de la Vie, de la Terre et de l'Environnement or SVTE). The course is part of a “common trunk” program that is compulsory for all Biology students. Subjects taught in the spring semester are: Professional Project (1.5 ECTS, European Credit Transfer and accumulation System), Zoology (3 ECTS), Plant Biology (3 ECTS), Cell Biology (3 ECTS),

Biochemistry-Chemistry (6 ECTS), Mathematics-Statistics (3 ECTS), Physics-Fluidics (3 ECTS), Informatics & Internet (3 ECTS), Functional Biochemistry (3 ECTS), Geology (3 ECTS) and English (3 ECTS). None of these courses had undergone important changes in teaching methods, nearly all science courses were taught in a (*low-structure*) transmission style throughout the study period. Exceptions were “Professional Project”, where students have to interview company-, teaching- or laboratory staff in order to learn about professional opportunities, and the English language course. All sciences courses had a comprehensive final exam, and from 2013 onwards, an additional mid-course exam. Both took place in dedicated exam weeks.

The introductory Cell Biology course amounted to 3 ECTS in the European Credit Transfer and accumulation System. This comprises 24 h face-to-face instruction, with 12 lecture slots of 1h20, 3 laboratory classes of 2h40 and an estimated 36 h of self-study. The course typically covers common Cell Biology topics. The lectures were given once a week and the three laboratory classes were dispersed over a period of 14 weeks (see figure 1). The laboratory classes mainly dealt with histology slides and tissue composition. The course was taught by a pool of 6 lecture teachers and 10 laboratory class teachers, who according to their availability and depending on the enrolment numbers, participated in the course for a certain number of hours. Lecture theatres had maximally 150 students, meaning that with an enrollment of 400, the cohort was separated over 3 lecture theatres and the course was repeated 3 times. Practical classes were limited to 32 students, accompanied by 2 teachers, meaning that with the same enrollment the class was repeated 13 times. The subjects of the in-course tests and exams were set prior to the course so that all teachers would make sure they covered the subject similarly and sufficiently. The first author, IK, participated in all courses, both in lectures and laboratory classes and was responsible for the course organization.

With respect to control of other possible variables, we controlled our study for exam or test equivalence by the use of similar exam formats, a combination of multiple choice questions and the annotation of images of subcellular compartments, and by giving students prior access to all possible exam questions. We could not control for instructor effects as this was impossible because of the change in student numbers and the resulting variable mobilization of teacher staff. Moreover, the pool of teachers had changed over the two teaching periods but turnover was independent of the teaching method employed (there was no choice or selection of teachers based on the course format). For these reasons we looked at student achievement over longer periods to

level out possible bias of teacher contributions. This was also the main argument why we returned to *low-structure* in 2017.

Experimental Intervention.

It concerns an experimental manipulation without randomization (American Psychology Association, 2020). The experimental intervention is an “instructional intervention” and comprises the implementation of a *moderate-structure* course format for an introductory Cell Biology course. The control situation is a *low-structure* format of the same course. We separate our analysis of the instructional intervention in two sets: one set of *low-structure* courses, years 2004, 2005, 2006 and 2017 and one set of *moderate-structure* courses years 2013, 2014, 2015 and 2016. We had returned to a *low-structure* course in 2017 in order to verify if possible change in student results were not a consequence of a structurally altered student population or teacher-related changes. Given the constraints imposed, we qualify this alternate schedule as "as good as possible."

Essential characteristics of the *moderate-structure* course format are: group-based exercises in class, group-based clicker quizzes (graded and non-graded) and more in-course knowledge assessments with a concomitant reduction in the length and weighting of the comprehensive final exam. Moreover, in the *moderate-structure* courses all course material, as well as interactive self-evaluation quizzes, were online (in French). For more detail, see table 1 below. We employ the term *moderate-structure* because our courses were devoid of weekly work assignments (no out-of-course activities) which according to Eddy & Hogan (2014) is an additional characteristic of a *high-structure* course.

Measures

We used in-course assessment scores, final-exam scores (see figure 1), Cell Biology-course grades and semester grade-point averages, as an instrument to study the impact of the course format on student achievement. As mentioned in the introduction, we make an important distinction between the regulation requirements for succeeding in-course knowledge assessments and the comprehensive final exam which occurred at the end of semester (weeks after course activities had ceased) (see figure 1). This is the reason we have analyzed course performance at two levels: 1) the graded in-course activities (tests, clicker quizzes and the mid-course exam) and 2) the comprehensive final exam (end of semester). In addition to this, we have also calculated

Table 1 *Low Structure (Conventional) and Moderate Structure (Experimental) Course Formats of the Cell Biology Courses that Are the Subject of this Study.*

Elements of the course / year	2004	2005	2006	2017	2013	2014	2015	2016
	low-structure				moderate-structure			
level of in-course grade points/structure	set 1		set 2		set 3		set 4	
Number of participating students	388	353	358	476	359	421	465	493
Socratic lecturing	x	x	x	x				
Lecture handout paper version	x	x	x	x				
On-line knowledge evaluation tool					x	x	x	x
On-line multimedia documents					x	x	x	x
On-line course slides	x	x	x	x	x	x	x	x
In-class videos					x	x	x	x
Non-graded group-based exercises in class					x	x	x	x
Non-graded group-based clicker quizzes in class					x	x	x	x
Graded group-based clicker quizzes in class (weighting of clicker score) ^a							x (0.25)	x (0.25)
Short in-course test (20 min) (weighting of each test score)	2 (0.125)	2 (0.125)						
Medium in-course test (40 min) (weighting of test score)			1 (0.25)	1 (0.25)	1 (0.25)	1 (0.25)		
Mid-course exam (1.5h) (weighting of mid-course exam score)					1 (0.35)	1 (0.35)	1 (0.35)	1 (0.35)
weighting of sum of in-course grades ^a	0.25	0.25	0.25	0.25	0.6	0.6	0.6	0.6
Final comprehensive exam (duration of exam)	1 (3h)	1 (3h)	1 (3h)	1 (3h)	1 (1.5h)	1 (1.5h)	1 (1.5h)	1 (1.5h)
weighting of final exam score ^a	0.75	0.75	0.75	0.75	0.4	0.4	0.4	0.4

Note: ^a the weighting represents the fraction of the total number of course points allocated to the course activity in question. It determines the impact on the total course grade. Example: the sum of points of the graded clicker-questions contributes to 25% of the course grade. The online course content: http://ressources.unisciel.fr/biocell/menu/co/module_menu_1.html

the Cell Biology pass rate for the different years (student retention rate). Pass rates is another parameter for determining course performance. Student retention is measured as the percentage of students with an overall course grade equal to or greater than 100/200 points.

Statistics and Data Analysis

The statistical analyses were performed with the 3.5.1 version of the R software (R Core Team, 2018), completed with the "car" 3.5.1 and the "mosaic" 3.5.3 packages.

Comparison of Overall Course Performance Corrected for Year Factor

As we deal with comparison of different student cohorts, in order to distinguish whether differences in were due to the effect of the *moderate-structure* (instructional intervention) or to differences in students' overall academic ability, we employed an analysis of variance for hierarchized plans ("nested factors ANOVA") where the "year" factor (cohort ability as represented by the semester grade-point average) is included in the instructional intervention factor (*low-* versus *moderate-structure* courses). We took the point that the students' average grade for all compulsory courses would represent their global academic ability (equivalent of a GPA) and therefore could serve as a means to normalize the variation between cohorts. The Cell Biology grade participates with a weighting of 10%. We have employed this detailed analysis for the two sets: years 2004-2006 and 2017, being representative of typical conventional, *low-structure*, courses and years 2013-2016, being representative of *moderate-structure* courses.

Comparison of In-course and Exam Scores of Low- and Moderate-structure Courses

We performed an ANOVA to see whether there were significant differences between the two sets: we compared *low-structure* (1) in-course, (2) final exam scores with *moderate structure* (3) in-course and (4) final exam scores. An ad hoc Tukey HSD test was performed to gain more insight in comparisons between the groups.

Comparison of Score Profiles in Low- and Moderate-structure Courses

We analyzed score differences between in-course assessments and the final exam by calculating the delta values (final exam minus in-course score). These values were plotted against the corresponding centered-course grade, which is the Cell Biology course grade minus 100 points. Course grades consist of 200 points of which 100 or more must be obtained to pass. This representation of student scores gives rise to four different classes: students who pass (+) or fail

(-) with either a higher exam scores than in-course scores (+) or the other way round (-) (see figure 3). We employed the chi-square test to compare the frequency distribution of the 4 classes observed in the *moderate-structure* set to that of the *low-structure* set, the latter we considered as the theoretical reference distribution. To refine the results of this global test, we performed post hoc binomial tests to identify the class or classes whose frequencies deviated significantly from the expected theoretical values.

Results

Moderate-structure Significantly Enhances Course Performance

We measured course performance and course retention relative to the grade-point average of all other (compulsory) disciplines taught in the corresponding semester and compared *low-* versus *moderate-structure* courses. When controlled for cohort ability we observe that *moderate-structure* courses had over a period of 4 years 8% higher grades (16 grade points/200) and 5 % higher retention rates compared to the *low-structure* courses (table 2). To analyze in more detail the impact of the instructional intervention relative to cohort academic ability we plotted for each

Table 2 *Learning Outcomes and Student Retention in Low- and Moderate-structure Courses*

	course format	retention	mean (SD)	F	Dfn	Dfd	η^2 ES	p	index (SD)
cell biology grade	low-structure (4 years)	63%	105.00 (36.9)	562.6	1	3312	0.1367 (13.7%)	< .0001	5.0 (18.0)
	moderate- structure (4 years)	82.5%	126.87 (32.8)						20.7 (20.8)
semester grade average	period “low”	53%	100.00 (23.6)	39.9	6	3312	0.0582 (5.8%)	< .0001	
	period “moderate”	67.5%	106.13 (22.0)						

Note: The values represent the grades and retention obtained by students over a four-year period in low and moderate structure courses. Grades are on a 200-point scale. Retention is expressed as the percentage of students who passed the course or passed the semester (requiring an average grade of 100 or higher). The relative contribution of the instructional intervention was analyzed with a nested ANOVA. A portion of the increase in course performance is explained by an overall improvement of cohort ability (semester grade-point average) and a large portion is explained by the instructional intervention (moderate-structure course format).

student the Cell Biology course grade against the corresponding semester grade-point average for the set of *low-structure* (years 2004, 2005, 2006 and 2017) and *moderate-structure* courses (years 2013-2016) (figure 2). In the *low-structure* performance in the Cell Biology course differed only slightly from their semester grade-point average with the regression line intercepting the y-axis at near zero. In contrast, the *moderate-structure* format revealed much higher Cell Biology course grades compared to corresponding semester grade-point average with the regression line intercepting the Y-axis at 30.7. We aligned individual Cell Biology grades with corresponding semester grades and calculated the differences between the two (“course index”). The nested-factor ANOVA shows that this “course index” is significantly lower in the *low-structure* compared to *moderate-structure* course ($M = 5.0$, $SD = 18.0$ versus $M = 20.7$, $SD = 20.8$ respectively, with $p < .0001$ (last column, table 2). The analysis did also reveal significant

Figure 2 Scatter Plot of Students’ Semester Grade Point Average Set Against their Cell Biology Course Grade.

Note: Two course formats are compared: in the left panel low-structure courses (years 2004-2006 and 2017) and in the right moderate-structure courses (years 2013-2016). The grey shaded area has been added as a reference; points above signify that students performed better in Cell Biology, compared to all other compulsory disciplines taught in the same semester. The dotted regression line in each graph is determined by least square regression calculations. Note that both the intercept with the y-axis and the slope of the lines vary between the two course-formats. Nearly all students benefit from moderate-structure but students with low semester grade-point averages benefit more than those with high semester grade-point averages (more equity).

variations in cohort abilities ($p < .0001$ (semester grade-point average). However, over a period of 4 years, “ability” explains only 5.8% (η^2 ES value in %) of grade variance whereas instructional intervention explains 13.7% (dominant). We take this to mean that we measure a real positive impact of the *moderate-structure* course format.

Low Achieving Students Benefit more from *Moderate-structure* Courses

Besides increased performance, the analysis of variance of the linear model revealed an interaction between the instructional intervention and the “student ability” ($F=43.9$, $Dfn=1$, $Dfd=3312$, $p < .0001$). This demonstrates that the effect size of *moderate-structure* courses is a function of global academic ability of the students. We therefore analyzed data of the two student populations separately (using the data plotted in figure 2). The regression line intercept with the y-axis is significantly higher than 0 in the *moderate-structure* courses ($b=30.7$, $p<0.0001$, $t=17.3$) whereas it is not in the *low-structure* courses. Moreover, the slope of the regression line of the *moderate-structure* courses is also significantly different from the one of the *low-structure* courses ($a_{mod} = 0.91$, $a_{low} = 1.05$) (figure 2). This reduction in slope and the high intercept value indicate that whereas all students benefit from *moderate-structure* courses, low-achieving students benefit more. We note that the reduction in slope is not a consequence of the fact that high achieving students can hardly improve their grade as they reach the 200 points (ceiling effect). The same slope is obtained when we exclude students with a semester grade average above 150.

A Significant Shift in Score Profiles between In-course and Final Exam Grades in Low-versus Moderate-structure Courses

Having established that student achievement benefits from a *moderate-structure*, we next wanted to identify where the gains were made. The in-course assessments and final exam scores shown in table 3 indicate that the in-course scores are always better than those on the final exam both in *low-* and *moderate-structure* formats. However, when we compare the results between the two formats we see a much larger gain in the in-course results of *moderate-structure* courses relative to the gain of final exam scores. An ANOVA and post hoc Tukey HSD test revealed that all the scores between the groups (1-4) are significantly different ($p < .001$) with the exception of group 1) in-course (*low-structure*) and 4) final exam (*moderate structure*) ($p=.598$). We conclude that the major gain of *moderate-structure* courses is made in in-course tests and not in the final exams.

Table 3 *In-course Tests and Final Exam Scores of Low- and Moderate-structure Courses.*

	low-structure		moderate structure		
	1) in-course	2) final exam	3) in-course	4) final exam	
N° of students	1491	1517	1657	1651	
Mean	116.2	107.7	146	115	
SD	38.5	38.9	31.5	43.8	
Source	DF	Sum of Square	Mean square	F statistic	P-value
Groups	3	1396194	465398	315.4	<.001
Error	6312	9313968	1475		
Total	6315	10710163	1695		

Moderate-structure Enhances the Frequency of Students who Underperform on the final Exam

In order to explain the differences in gain of in-course and final exam scores we looked at the differences in the grades of individual students. For each student, we calculated the difference between final exam and in-course grade (x-axis) and plotted this difference against a centered course grade (course grade reduced by 100 points, which is the minimum number of points needed to pass). In doing so we obtained a profile with 4 classes: (+/-), students who passed the course with an exam score inferior to in-course scores; (+/+) students who passed the course with an exam score superior to in-course scores; (-/-), students who failed the course with an exam score inferior to in-course scores; and (-/+), students who failed the course with an exam score superior to in-course scores. The percentages of students in each of the classes is shown in figure 3. There are two important changes. Firstly, as expected, an important shift to higher grades in *moderate-structure* courses. The percentage of failing students had reduced from 33.4 to 14.4%. Secondly, whereas in *low-structure* we observe a near equal distribution of positive (exam-score better than in-course) and negative difference (exam-score worse than in-course) ($\bar{x} \text{ diff} = 0, SD = 29$), in *moderate structure* we see a near depletion of positive differences in favor of negative differences ($\bar{x} \text{ diff} = -36, SD = 22$). The number of students who performed better in the final exam was reduced from 46.2 to 9.6%. The vast majority of students in *moderate-structure* courses are now in the class of achieving students with a final exam score inferior to in-course scores (+/-, 77%). The Pearson's Chi-squared test revealed highly significant differences

between the frequencies of the 4 classes (X-squared = 1318.5, $df = 3$, $p < 2.2e-16$). Subsequent exact binomial tests revealed that the frequencies of all 4 classes are significantly different between *low*- and *moderate-structure* courses ($p < .0001$).

Figure 3 Scatter Plot of Score Differences Between Exam and In-course Assessments of Individual Students Set Against their Centered Cell Biology Course Grade (Grade-100 points).

Note: The difference between the results of in-course tests and those of the final exam (x-axis) is plotted against a centered course grade (y-axis). The zero value on the y-axis corresponds to 100 points, the minimal necessary grade to pass the course. The scatterplot shows a sharp reduction in the number of failing students (lower section) but also a significant shift in the orientation of the delta values between in-course and exam scores. An overwhelming majority of students now perform worse in the exam than in the in-course tests.

The distribution of Scores Reveals a Bimodal Distribution for Exam Scores in Moderate-structure Courses.

After identifying an enhanced frequency of students who underperform on the exam, we next looked how this is reflected in the distribution of scores (non-paired in-course and final-exam scores). We have chosen 10 bins per standard deviation for sufficient resolution which practically translated into bins of 5 points on a scale of 200 (Figure 4). The histograms show that for *low-structure* courses, the distribution of scores differs little between the in-course test and the final exam; both seem to follow an approximatively normal distribution, with a small tail of very low scores. The entire population seems to have moved to slightly lower scores on the final

exam (shift to the right). The situation is quite different in courses with *moderate-structure*, where we get non-overlapping distributions of scores. While the in-course test shows a highly skewed and compact unimodal distribution with very few low scores, as if dealing with a homogeneous successful student population, the distribution of final-exam scores smears across the entire range of scores, indicative of highly heterogeneous population. However, the distribution of exam scores differs between the two course formats in that it appears as a bimodal distribution in *moderate-structure* as a result of an increased number (+177) of very high scores (in the range of 150-200 points) and an increased number (+67) of very low scores (in the range of 0-60 points). The score distribution shown in figure 4 combined with the data shown in figure 3 point to the existence of a subpopulation of students that more or less maintains initial gains (persists) and a subpopulation that absolutely fails to maintain those gains. (See also Figure S1, in the Supplementary Material, for the score distributions across individual years. The histograms show that the difference between low- and moderate structure holds for each year).

Figure 4 *Distribution of Scores Between In-course and Final Exam Grades of Low- and Moderate-structure Course Formats.*

Discussion

We show that, when corrected for cohort capability, the *moderate-structure* courses significantly improved student performance and retention, with on average 8% higher grades and 5 % more pass scores when corrected for academic ability of the cohorts. Indeed, in particular, students with a low semester grade-point average, benefit from course structure. We can refute our primary hypothesis (H0) and argue that increased structure does affect study outcomes significantly and positively. Our study confirms findings of numerous other studies (Crouch &

Mazur, 2001; Ebert-May et al., 1997; Eddy & Hogan, 2014; Freeman et al., 2007, 2011, 2014; Haak et al., 2011; Knight & Wood, 2005; Tanner, 2013; Wood, 2003). We interpret our results to mean that the instructional intervention (*moderate structure*) is sufficiently robust to deliver measurable improvements under the variable conditions prevailing in large open-enrollment universities in France. With respect to the second research question, whether or not all students are moving toward a higher level of internalization of learning and that this is reflected in better achievement with a similar distribution of scores of all knowledge assessments, we have to refute our primary hypothesis (H0). The distribution of scores reveal that while nearly all students perform very well on in-course assessments in *moderate-structure* courses, a large subpopulation of students fails to persist their achievement and performed poorly on the final exam. As a consequence we observe a large and significant shift in the frequency of student that score worse in the final exam than in the in-course assessments (from 53.6 to 90.3%). From the point of view of Cell Biology knowledge, the lack of persistence (abundant failures in the final exam) means that there is still a similar number of students (comparable to *low-structure*) who have not processed a substantial part of the course content at all and will carry this deficit into the next year. Below we discuss possible reasons for the lack of persistence in student performance in *moderate-structure* courses. We focus on persistence because it is a critical factor in academic achievement (Duckworth & Seligman, 2005).

Students Calculate a Minimal Necessary Effort or Lack Incentive for the Final Exam

At first glance, the important shift in the profile of in-course versus final-exam scores (figure 3) and the shift in score distribution (figure 4) that we observe in *moderate-structure* courses gives the impression that some kind of self-limitation is operating. A certain amount of self-limitation is not so bad in itself, as long as the final exam grades remain adequate. But they are not for a large number of students. Some 36% of students fail the final exam in *moderate-structure*, compared to 33% in *low-structure*. A much heard argument is that numerous students deliberately “calculate” their learning effort in order to pass the course with minimal input. In the case of dedicated exam weeks, this allows certain students to spend more time preparing for the final exams of other disciplines for which they do not have good in-course grades. Alternatively, colleagues argue that the high grades obtained during the course took away the incentive to prepare well for the final exam. We do not exclude these possibilities but if calculated input and/or lack of incentive determined the learning effort for the final exam, we should over the course of the eight years identify an inverse relationship between annual mean in-course and

mean final-exam scores and we do not ($R^2 = 0.090$, data not shown). Moreover, if students could calculate minimal effort, why do so many fail the entire course because their exam score was much too low to be compensated by in-course scores (some 7% of the total student population in moderate-structure courses)? As a last point, we note that 70% of the students who had not passed the final exam had enrolled for the second year. Thus, many of the insufficient final-exam scores also cannot be explained by the fact that students had already decided to leave the study.

The Final Comprehensive Exam Remains a Destructive Friction for a Subpopulation of Students

Another possible explanation for final-exam failure is that, despite the gradual reduction of teacher control over learning, it remains a destructive friction for a good number of students. The organization of all final exams in a dedicated exam week, which is the practice in the university where this study was conducted, exacerbates this friction. Such (unnecessary) concentration of exams increases the apprehension of assessment, it prompts superficial learning and leads students to misjudge the workload. A possible solution is to abolish special exam weeks and only hold in-course knowledge tests. However, there are reasons to appreciate the role of a comprehensive final exam when combined with in-course testing. A comprehensive final exam can help students gain a more integrated view of the subject and, importantly, the second repetition of the course material reinforces long-term knowledge retention (Carpenter et al., 2012; Hartwig & Malain, 2022; Roediger & Butler, 2011). Students will have more ready knowledge at the start of the next year, allowing them to better integrate new information. Repetition of course material could, of course, be built into repeated in-course tests, but that would require more instruction-free moments in the semester to allow students to review course content. Alternatively, comprehensive final exams are offered at different moments and students decide when they are ready for the test.

Ego Depletion as a Cause of Failing Persistence?

The destructive friction may be caused by cognitive failure, too much knowledge assimilation (high cognitive load) in too short a time (Merriënboer & Sweller, 2005). Again, this cognitive load was compounded by the fact that the final exam took place during a dedicated exam week so that students had to juggle a number of disciplines at once. This further complicated matters because it required a very high level of internal guidance (self-regulation) (Ten Cate et al, 2004; Zimmerman, 2013). This is not an entire satisfactory explanation because

although the mid-course exam produces less destructive friction (Vermunt & Verloop, 1999), the conflict with other disciplines was not less and yet the results were overwhelmingly good. Moreover, students had a very precise idea of what was needed to pass and they had access to all possible exam questions. The final exam was not an exploration of unknown territory, but a mere repetition of a similar exercise they knew they good do well. At least for Cell Biology, most students should have had a good sense of self-efficacy (Bandura, 1977) by the end of the lectures.

In addition to cognitive failure, destructive friction in preparation for the final exam can be amplified by a depletion of energy, also known as ego depletion, of some of the students (Baumeister & Vohs, 2007; Ryan & Deci, 2008). Baumeister et al., (p. 1253, 1998) defined ego depletion as “a temporary reduction in the self’s capacity or willingness to engage in volitional action (including controlling the environment, controlling oneself, making choices, and initiating action), caused by prior exercise of volition”. Supporting this view, experiments have shown that behaviors involving self-controlling actions deplete energy, as manifest in decrements in performance or persistence at subsequent tasks (Baumeister & Vohs, 2007). A key point here is that the degree of energy depletion depends on the cost of learning and that this cost is not the same for all learners (Ryan & Deci, 2008). Over time and depending on learning conditions, some students will deplete more than others. It may also be the case that students do not have similar starting levels and that over time some students may not have enough energy to successfully complete the final exam. If we place the task value, namely mastering the subject of cell biology, in a model that describes the utility of a task (U) as the sum of its value (R) minus its cost (C), written as $U=R-C$ (Jara-Ettinger, et al., 2016), it follows that students for whom learning costs a lot, resulting in a high degree of ego depletion, are more likely to lose a sense of utility. Alternatively, students with a lower perceived value of the task (Eccles & Wigfield, 2002), at similar learning costs, may begin with a lower energy level and therefore lose a sense of utility more quickly as they progress through the course. Both the increased engagement during the course (participating in repeated knowledge assessments) and the abrupt transition from an external/shared-guidance to internal-guidance of learning (high self-regulation) in the period leading up to the exam may have been too much for some students. Ego depletion may be a factor, in addition to possible inappropriate teacher skills (Andrews et al., 2011), that explains why courses that emphasize active learning do not always lead to overall better performance. We are currently examining how students interact with essential learning documents throughout the course. Do they maintain learning activities outside the lectures or do we see a reduction in

consulting those documents and how does that correlate with in-course and exam scores (a learning analytics approach)?

Conclusions

The introduction of courses with a *moderate structure* is associated with a significant increase in course performance. This gain is mainly determined by extraordinarily good grades on in-course assessments, as if we were dealing with an almost homogeneous successful student population. However, the final exam results show a very different picture, one that is consistent with the highly heterogeneous population we repeatedly observe in *low-structure* courses. The initially good results do not last; the final exam remains a destructive friction for a good portion of students. As a result of the still many poor final-exam results, we retain a large group of students who have not adequately assimilated a significant portion of the necessary cell biology knowledge and enter the second year with significant gaps.

Shortcomings of the Study

The study has a number of shortcomings which should be considered in interpreting the findings. The first deals with the lack of internal control (student and teacher population), the second with the particular final-exam conditions and the third with the open enrolment student population. Selective (elite) universities may not face the same problem, or to a lesser extent and institutions that do not impose a comprehensive final exam may see better outcomes. The lack of persistence in Cell Biology may be due to a growing frustration of student interest because of the *low structure* of courses in all other science subjects. Finally, we provide no evidence as to why students fail to maintain their initial high scores and are only offering suggestions.

Acknowledgements.

Gerard Tramu, one of the Cell Biology core teachers, and Pascale Rallion and Paul Bensamoun, from the teaching support department (MAPI) of the University of Bordeaux, are acknowledged for their important help with the development of the multimedia resources and interactive questions. We thank the numerous Biology teachers of the University of Bordeaux who participated in the course. We are grateful to UniSciel (the online Science University of France (www.unisciel.fr) for the hosting the [Cell Biology resources](#) and make them accessible to a very large public.

Conflict of Interest Statement:

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- American Psychology Association. (2020). *Publication Manual of the American Psychological Association* (7th ed.). <https://doi.org/10.1037/00000165-000>
- Andrews, T.M., Leonard, M.J., Colgrove, C.A. and Kalinowski, S.T. (2011). Active learning not associated with students learning in a random sample of college biology courses. *CBE Life Sciences Education*, 10, 394-505. <https://doi.org/10.1187/cbe.11-07-0061>
- Astin, A.W. (1972). College dropouts: a national profile. *ACE Research Reports* 7. Washington, D.C.: American Council on Education. <https://eric.ed.gov/?id=ED059691>
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Baumeister, R. F., & Vohs, K. D. (2007). Self-Regulation, ego depletion, and motivation. *Social and personality psychology compass*, 1(1), 115-128. <https://doi.org/10.1111/j.1751-9004.2007.00001.x>
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of personality and social psychology*, 74(5), 1252-1265. <https://doi.org/10.1037/0022-3514.74.5.1252>
- Beishuizen, J.J., & Stoutjesdijk, E.T. (1999). Study strategies in a computer assisted study environment. *Learning and Instruction*, 9(3), 281-301. [https://doi.org/10.1016/S0959-4752\(98\)00027-9](https://doi.org/10.1016/S0959-4752(98)00027-9)
- Briggs, A.R.J., Clark, J., & Hall, I. (2012): Building bridges: understanding student transition to university, *Quality in Higher Education*, 18(1), 3-21. <https://doi.org/10.1080/13538322.2011.614468>

- Brownell, S.E., & Tanner, K.D. (2012). Barriers to faculty pedagogical change: lack of training, time, incentives, and tension with professional identity. *CBE Life Sciences Education*, 11(4), 339-346. <https://doi.org/10.1187/cbe.12-09-0163>
- Caldwell, J.E. (2007). Clickers in the large classroom: current research and best-practice tips. *CBE-Life Sciences Education* 6, 9-20. <https://doi.org/10.1187/cbe.06-12-0205>
- Cameron, R.B., & Rideout, C.A. (2020). “It’s been a challenge finding new ways to learn”: first-year students’ perceptions of adapting to learning in a university environment. *Studies in Higher Education*, <https://doi.org/10.1080/03075079.2020.1783525>
- Carpenter, S.K., Cepeda, N.J., Rohrer, D., Kang, S.H.K., & Pashler, H. (2012) Using Spacing to Enhance Diverse Forms of Learning: Review of Recent Research and Implications for Instruction. *Educational Psychology Review* 24(3):369–378. <https://doi.org/10.1007/s10648-012-9205-z>
- Crouch, C.H., & Mazur, E. (2001). Peer instruction: ten years of experience and results. *American Journal of Physics*, 69, 970-977. <https://doi.org/10.1119/1.1374249>
- Duckworth, A. L., & Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological science*, 16(12), 939-944. <https://doi.org/10.1111/j.1467-9280.2005.01641.x>
- Ebert-May, D., Brewer, C., & Allred, S. (1997). Innovation in large lectures: teaching for active learning. *BioScience*, 47, 601-607.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109-132. <http://dx.doi.org/10.1146/annurev.psych.53.100901.135153>
- Eddy, S.L., & Hogan, K.A. (2014). Getting under the hood: how and for whom does increasing course structure work. *CBE-Life Sciences Education*, 13, 453-468. <https://doi.org/10.1187/cbe.14-03-0050>
- Ferrare, J.J. (2019). A multi-institutional analysis of instructional beliefs and practices in gateway courses to the sciences. *CBE-Life Sciences Education* 18:ar26, 1-16. <https://doi.org/10.1187/cbe.17-12-0257>

- Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor N., Jordt, H., & Wenderoth, M.P. (2014). Active learning increases student performance in science, engineering and mathematics. *PNAS (USA)*, *111*, 8410-8415. <https://doi.org/10.1073/pnas.1319030111>
- Freeman, S., Haak, D., Wenderoth, M.P. (2011). Increased course structure improves performance in introductory biology. *CBE-Life Sciences Education*, *10*, 175-186. <https://doi.org/10.1187/cbe.10-08-0105>
- Freeman, S., O'Connor, E., Parks, J.W., Cunningham, M., Hurley, D., Haak, D., Dirks, C. & Wenderoth, M.P. (2007). Prescribed active learning increases performance in introductory biology. *CBE-Life Sciences Education*, *6*, 32-139. <https://doi.org/10.1187/cbe.06-09-0194>
- Fryer, L.K., & Vermunt, J.D. (2018). Regulating approaches to learning: testing learning strategy convergences across a year at university. *British Journal of Educational Psychology*, *88*, 21-41. <https://doi.org/10.1080/02602930500099102>
- Gasiewski, J.A., Eagan M.K., Garcia G.A., Hurtado S. & Chang M.J. (2012). From gatekeeping to engagement: a multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Research in Higher Education*, *53*, 229-261. <https://doi.org/10.1007/s11162-011-9247-y>
- Haak, D.C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science*, *332*, 1213-1216. <https://doi.org/10.1126/science.1204820>
- Hartwig, M.K., & Malain, E.D. (2022). Do students space their course study? Those who do earn higher grades. *Learning and Instruction* *77*, 101538. <https://doi.org/10.1016/j.learninstruc.2021.101538>
- Harvey, L., Drew, S. & Smith, M. (2006). The First-Year Experience: A review of literature for the Higher Education Academy (London, *The Higher Education Academy*). (Accessed 14/10/2020): http://www.improvingthestudentexperience.com/library/UG_documents/first_year_experience_full_report_Harvey_et_al.pdf

- Hillman, K. (2005). The First Year Experience: The transition from secondary school to university and TAFE in Australia, Longitudinal Surveys of Australian Youth, Research Report 40, Australian Council for Educational Research and the Australian Government Department of Education, Science and Training.
https://research.acer.edu.au/lsay_research/44
- Jara-Ettinger, J, Gweon, H, Schulz, LE, & Tenenbaum, JB. (2016). The naïve utility calculus: computational principles underlying commonsense psychology. *Trends in Cognitive Sciences* 20, 589-604. <https://doi.org/10.1016/j.tics.2016.05.011>
- Knight, J.K., & Wood, W.B. (2005). Teaching more by lecturing less. *Cell Biology Education* 4, 298–310. <https://doi.org/10.1187/05-06-0082>
- Kramer, I.M., & Kusrkar, R.A. (2017). Science-writing in the blogosphere as a tool to promote autonomous motivation in education. *The Internet and Higher Education* 35, 48-62.
<http://dx.doi.org/10.1016/j.iheduc.2017.08.001>
- Merrienboer, J.J.G, & Sweller J. (2005). Cognitive load theory and complex learning: recent developments and future directions. *Educational Psychology Review* 17, 147-177.
<https://doi.org/10.1007/s10648-005-3951-0>
- Nelson, C.E. (2008). Teaching evolution (and all of biology) more effectively: strategies for engagement, critical reasoning, and confronting misconceptions. *Integrative and Comparative Biology* 48, 213–225. <https://doi.org/10.1093/icb/icn027>
- Panadero, E. (2017). A review of self-regulated learning: six models and four directions for research. *Frontiers Psychology* 8, 422. <https://doi.org/10.3389/fpsyg.2017.00422>.
- R Core Team. (2018). R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org/>
- Roediger, H.L., & Butler, A.C. (2011). The Critical Role of Retrieval Practice in Long-Term Retention. *Trends in Cognitive Sciences* 15(1), 20–27.
<https://doi.org/10.1016/j.tics.2010.09.003>

- Ryan, R.M., & Deci, E.L. (2000a). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist* 55(1), 68-78.
<https://doi.org/10.1037/0003-066X.55.1.68>
- Ryan, R.M., & Deci, E.L. (2000b). Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemporary Educational Psychology* 25, 54-67.
<https://doi.org/10.1006/ceps.1999.1020>
- Ryan, R. M., & Deci, E. L. (2008). From ego depletion to vitality: Theory and findings concerning the facilitation of energy available to the self. *Social and Personality Psychology Compass* 2(2), 702-717. <https://doi.org/10.1111/j.1751-9004.2008.00098.x>
- Ryan, R.M., & Deci, E.L. (2017). *Self-determination theory: basic psychological needs in motivation, development, and wellness*. The Guilford Press. ISBN 9781 4625 2876 9.
<https://doi.org/10.1521/978.14625/28806>
- Sander, P., Stevenson, K., King, M. and Coates, D. (2000). University students' expectations of teaching. *Studies in Higher Education* 25(3), 309–323.
<https://doi.org/10.1080/03075070050193433>
- Smith, M.K., Wood, W.B., Adams, W.K., Wieman, C., Knight, J.K., Guild, N., & Su, T.T. (2009). Why peer discussion improves student performance on in-class concept questions. *Science* 323(5910), 122-124. <https://doi.org/10.1126/science.1165919>
- Tanner, K.D. (2013). Structure matters: twenty-one teaching strategies to promote student engagement and cultivate classroom equity. *CBE-Life Sciences Education* 12(3), 322-331.
<https://doi.org/10.1187/cbe.13-06-0115>
- Ten Cate, O., Snell, L., Mann, K., & Vermunt, J. (2004). Orienting teaching toward the learning process. *Academic Medicine*, 3, 219-228. No doi.
- Tinto V. (1975). Dropout from higher education: a theoretical synthesis of recent research. *Reviews of Educational Research* 45(1), 89-125.
<https://doi.org/10.3102/00346543045001089>

- Townsend, M.A.R., & Hicks, L. (1997), Classroom goal structures, social satisfaction and the perceived value of academic tasks. *British Journal of Educational Psychology* 67, 1-12. <https://doi.org/10.1111/j.2044-8279.1997.tb01222.x>.
- Vermunt, DJ, & Verloop N. (1999). Congruence and friction between learning and teaching. *Learning and Instruction*, 9, 257-280. [https://doi.org/10.1016/S0959-4752\(98\)00028-0](https://doi.org/10.1016/S0959-4752(98)00028-0)).
- Wood, W. B. (2003). Inquiry-based undergraduate teaching in the life sciences at large research universities: A perspective on the Boyer Commission Report. *Journal of Cell Biology Education* 2(2), 112–116. <https://doi.org/10.1187/cbe.03-02-0004>
- Yorke, M., & Longden, B. (2008). The first-year experience in higher education in the UK: final report. *The Higher Education Academy*.
https://www.heacademy.ac.uk/system/files/FYEFinalReport_1.pdf (accessed 31 dec 2021)
- Zimmerman, B.J. (2013). From cognitive modeling to self-regulation: a social cognitive career path. *Educational Psychologist* 48(3), 135-147.
<https://doi.org/10.1080/00461520.2013.794676>

Figures & table captions

Figure 1

Time-line of Course Activities: Lectures, Laboratory Classes, In-class Clicker Quizzes, Tests and Exams.

Note: This is a summary time-line of all courses over an 8-year period. Exactly what activities took place in low- and moderate-structure courses is described in Table 1.

Figure 2

Scatter Plot of Students' Semester Grade-Point Average Set Against their Cell Biology Course Grade.

Note: Two course formats are compared: in the left panel low-structure courses (years 2004-2006 and 2017) and in the right moderate-structure courses (years 2013-2016). The grey shaded

area has been added as a reference; points above signify that students performed better in Cell Biology, compared to all compulsory disciplines taught in the same semester. The dotted regression line in each graph is determined by least square regression calculations. Note that both the intercept with the y-axis and the slope of the lines vary between the two course-formats. Students perform better in a moderate-structure course but those with low semester grade-point averages benefit more than those with high semester grade-point averages.

Figure 3

Scatter Plot of Score Differences Between Exam and In-course Assessments of Individual Students Set Against their Centered Cell Biology Course Grade (Grade-100 points).

Note: The difference between the results of in-course tests and those of the final exam (x-axis) is plotted against a centered course-grade (y-axis). The zero value on the y-axis corresponds to 100 points, the minimal necessary grade to pass. The scatterplot shows a sharp reduction, in the number of failing students (lower sections, from 33.4 to 14.4%) but also a significant shift in the orientation of the delta values between in-course and exam scores. An overwhelming majority (left sections, 90.3%) of students now perform worse in the exam than in the in-course tests.

Figure 4

Distribution of Scores Between In-course and Exam Grades of Low- and Moderate-structure Course Formats

Note: The distributions represent the scores obtained by students over a four-year period in courses with low- or moderate-structure. In the case of low-structure, the two distributions overlap with a slight shift to lower scores for the final exam. This corresponds to a slightly lower average final exam score. In the case of moderate-structure, the distributions barely overlap. The in-course grades are high and have relatively low dispersion. In contrast, the dispersion of the final exam results is huge and shows a larger percentage of very high and of very low grades, as if there were two subpopulations (tentatively indicated with dotted lines).

Table 1

Low Structure (Conventional) and Moderate Structure (Experimental) Course Formats of the Cell Biology Courses that Are the Subject of this Study.

Note: ^a the weighting represents the fraction of the total number of course points allocated to the course activity in question. It determines the impact on the total course grade. Example: the sum of points of the graded clicker-questions contributes to 25% of the course grade. The online course content:

http://ressources.unisciel.fr/biocell/menu/co/module_menu_1.html

Table 2

Learning Outcomes and Student Retention in Low- and Moderate-structure Courses.

Note: The distributions represent the scores obtained by students over a four-year period in low- and moderate-structure courses. Grades are on a 200-point scale. Retention is expressed as the percentage of students who passed the course or passed the semester (requiring a grade of 100 or higher). The relative contribution of the instructional intervention was analyzed with a nested ANOVA. With respect to grades, a portion of the increase in course performance is explained by an overall improvement of cohort ability (semester grade point average) and a large portion is explained by the instructional intervention (moderate-structure format).

Table 3

In-course tests and Final Exam Scores of Low- and Moderate-structure Courses.

Supplemental Material. Figure S1 Score distribution over the separate years.

Note the systematic shift in in-course- and exam-score distributions between low- and moderate-structure courses. Whereas distributions of in-course and final exam fully overlap in low-structure, they never do in moderate-structure courses. Note also that when we shift back from moderate to low-structure in 2017, the results are again similar to previous low-structure courses. We regard this as additional evidence that changes in course performance and score distribution are not the result of structural changes in the student population.