

Article

Declining Student Performance and Satisfaction following Back-to-Back Scheduling of Foundational Science Exams: Experience at a Large US Osteopathic Medical School

Martin Schmidt ^{1,*} , Brian Pinney ²  and Sarah Werning ³¹ Department of Biochemistry and Nutrition, Des Moines University, Des Moines, IA 50311, USA² Center for Educational Enhancement, Des Moines University, Des Moines, IA 50311, USA;

brian.pinney@dmu.edu

³ Department of Anatomy, Des Moines University, Des Moines, IA 50311, USA; sarah.werning@dmu.edu

* Correspondence: martin.schmidt@dmu.edu

Abstract: Examinations are a significant source of academic stress, particularly in the demanding environment of medical education. In order to reduce the burden of anxiety, burnout and depression among students, medical schools aim at reducing academic stress by exploring alternatives to frequent, high-stakes assessment schedules. The bundling of examinations into integrated block assessments has emerged as a successful strategy to reach this goal, as the resulting reduction in examination days can provide uninterrupted periods of study time that allows for a deeper understanding of material as well as an opportunity for students to take wellness breaks between exams. The present study examines the outcomes of a natural experiment on back-to-back exam scheduling in two related medical school courses post hoc. The outcomes of the scheduling change on the academic performance and satisfaction of examinees were analyzed with a mixed-methods approach. The data show that the transition from a spaced-out to a back-to-back exam schedule was accompanied by a drop in academic performance and student satisfaction with the curricular schedule. The data presented suggest that without proper curricular integration, the block scheduling of exams has negative effects on learning outcomes and student satisfaction.

Keywords: exam scheduling effects; undergraduate medical education; assessment strategies



Citation: Schmidt, M.; Pinney, B.; Werning, S. Declining Student Performance and Satisfaction following Back-to-Back Scheduling of Foundational Science Exams: Experience at a Large US Osteopathic Medical School. *Educ. Sci.* **2022**, *12*, 94. <https://doi.org/10.3390/educsci12020094>

Academic Editor: Marius Boboc

Received: 2 December 2021

Accepted: 28 January 2022

Published: 29 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The assessment of learning outcomes is an essential element of curriculum design, but even well-designed assessments are a major source of stress and anxiety in University students [1]. While the methods and scheduling of assessments have a significant influence on the measured performance of examinees [2], advanced assessment strategies such as integrated block testing go beyond a mere improvement in measurement, and instead improve the actual learning outcomes of the curriculum [3,4]. Particularly in the demanding medical school curricula, with frequent high-stakes examinations, assessment formats and schedules are a major source of student stress [5], and need to be carefully designed to provide an authentic measure of learning, as well as to favor the understanding of key concepts over the short-term memorization of facts [6].

While it seems obvious that the scheduling of exams can have a major impact on measured performance and student satisfaction, only a few studies have attempted to quantify scheduling effects on academic performance in isolation of any other curricular changes [7,8]. In particular, the influence of time between exams on academic performance has been understudied, and the only available comprehensive data on such scheduling effects are limited to an analysis of the performance of high school students [2]. In the medical school setting, changes in exam schedule are usually part of a global change in assessment strategy [9]. Systems-integrated scheduling (SIS) is a well-established curricular

approach to foster the integration of subject material through coordinated teaching and assessment across disciplines in a coordinated block format [10]. While the change from discipline-specific to integrated exams was described as broadly positive [11], a detailed quantitative analysis showed that exam integration in SIS curricula has a more complex and not entirely positive impact on student academic performance [12], with students scoring lower overall on integrated exams than on the more traditional separate discipline-specific exams, but showing improved performance on the more complex questions of higher Blooms taxonomic classification. As was concluded at the time of implementation, SIS requires deliberate careful planning to leverage the benefits of content integration for more meaningful learning [10].

The present study was primarily undertaken to assess the effects of the back-to-back (B2B) scheduling of two basic science course exams on student academic performance in the preclinical curriculum at a large US osteopathic medical school. No intentional effort was given within this study to integrate course content or assessment items. A second, related aim of the present study was to capture student opinions on the B2B testing schedule through a systematic analysis of comments on the annual student satisfaction survey. In its narrow scope to assess scheduling effects in isolation of any other curricular changes, the present study differs significantly from related work on the effects of SIS on learning outcomes. While this difference limits the applicability of the extensive theoretical framework on curricular integration [13], it highlights the unique value of the study as a unique opportunity to disentangle the effects of exam scheduling from other effects of curricular integration.

In 2019, examinations in Biochemistry/Genetics and Medical Cell & Tissue Biology courses were administered on the same day with a 15 min break, instead of being spaced 7–14 days apart as had been done in prior years. It was reasoned that even in the absence of other curricular changes, the simultaneous testing of related subjects would contribute to a deeper understanding of complex materials by encouraging integrated study, and would help to improve student satisfaction by reducing the number of exam days. Based on this rationale, it was hypothesized that the B2B schedule would lead to an increase in measured learning outcomes, accompanied with an increase in the subjective measure of satisfaction with the curriculum—two measures that are long known to be correlated [14,15]. In addition, it was hypothesized that the increased length of the combined assessments would be recognized as beneficial, as it would allow students to build stamina for medical licensing examinations and could give insights into potential effects of exam length on performance in the medical school setting [16,17].

2. Materials and Methods

2.1. Experimental Design and Intervention

In the fall semester of 2019, the Des Moines University College of Osteopathic Medicine (DMU-COM), a large US Osteopathic Medical School, changed the scheduling of exams in two parallel courses, Biochemistry/Molecular Genetics (BMG, 4 credits) and Medical Cell & Tissue Biology (MCTB, 4 credits). Following an administrative decision with limited faculty and student input, the exam schedule was changed from a spaced-out system (SO) to a back-to-back model (B2B; see Figure 1). No other curricular changes were made, and no significant changes were made to the exams. In both models, students were given 75–83 min to complete 50–55-question multiple choice examinations. In the 2018 SO model, the exams were 7–14 days apart, while in the 2019 B2B model the MCTB exam was scheduled following the BMG exam after a 15 min break (necessary for technical reasons). Students who submitted the first exam before the end of the allotted time were thus given a longer break between exams.

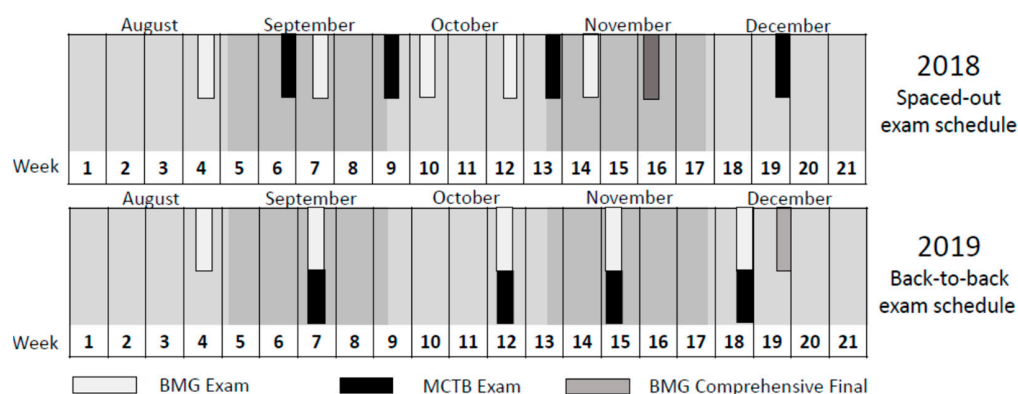


Figure 1. Comparison of the spaced-out and back-to-back testing schedules for Biochemistry/Molecular Genetics (BMG) and Medical Cell & Tissue Biology (MCTB) employed in this study. In 2018, first-year students in the osteopathic and podiatric medicine programs at Des Moines University (DO22 and DPM22) were assessed on separate days. In 2019, the BMG and MCTB course exams were administered back-to-back with a 15 min break in between. Note that the scheduling of the first BMG exam was not changed, meaning that the assessment could serve as an internal control of cohort academic performance.

2.2. Assessment Strategies and Exam Administration

DMU-COM uses a competency-based assessment strategy featuring both formative and summative examinations. Course content and exam questions are linked to the Osteopathic Core Competencies [18], with competencies II. Medical Knowledge and III. Patient Care being most frequently taught and assessed. For each MCTB and BMG exam, students were given a full-length formative practice quiz, which is not included in the grade. As an additional formative element in both exams and practice quizzes, students are shown the question rationales and linked course objectives after the exam has concluded. Exams were administered on-campus with live proctoring using the Examsoft testing platform (Examsoft, Farmers Branch, TX, USA). Students were given a presentation during orientation to explain the nature of and the rationale behind the scheduling change. Study authors (MS and SW) served as directors for the BMG and MCTB courses, respectively. In addition to these courses, students were also required to take courses in Anatomy (5 credits), and other program-specific offerings such as Clinical Medicine (1.5 credits) and Osteopathic Manual Medicine (2.5 credits). The exams in all other courses followed the SO schedule, and did not coincide with the combined BMG/MCTB tests.

This curricular change set up a natural experiment, which allowed a post-hoc cohort study comparing measures of academic performance and student satisfaction between the 2018 cohorts of 271 first-year osteopathic (DO22) and podiatric medicine (DPM22) students (assessed with SO schedule) and the 270 members of the respective 2019 cohorts (DO23 and DPM23, assessed with B2B schedule).

2.3. Characterization of Study Cohorts

De-identified academic and biographical data of the DO22, DO23, DPM 22 and DPM23 study cohorts were obtained from the University Admissions Department and compared for significance using two-tailed students' *t*-test for all values except for male/female ratio, which was analyzed for significance with a Chi Square test (Microsoft Excel, Redmond, WA, USA).

2.4. Measurements of Academic Performance

The influence of the exam schedule on students' academic performance was determined by analysis of exam item performance (difficulty and point biserial), exam reliability (KR20 values) and course grades using the tools available through the exam administration software. Items that were changed or replaced during the 2018/2019 transition were excluded before further analysis. Exam reliability, average point biserial and the

difficulty of 10 BMG/MCTB exams, as well as course grades, were analyzed for significance in differences using non-parametric testing (Mann–Whitney U testing; SPSS, Chicago, IL, USA).

2.5. Measures of Student Satisfaction

Student satisfaction with the curricular scheduling was determined by qualitative analysis of responses to the routine annual spring survey of students of osteopathic medicine. This comprehensive survey is conducted annually by the COM Dean’s office and collects information to support improvements in educational programming and student learning. Survey participation is optional, and the survey contains no forced items. The OMS1/OMS2 survey contains 100 response items, including 3 free-text response items. Survey participation among 2018 and 2019 first- and second-year students was between 61% and 67%. Free-text responses relating to the exam schedule were coded by two investigators (BP, MS) for positive, neutral or negative impressions on the B2B testing schedule, for emerging common themes describing the problem with the B2B schedule, as well as for statements indicating a connection between exam schedule and stress. Inter-rater reliability testing was performed using Cohen’s kappa [19] (Microsoft Excel, Redmond, WA, USA). Text narratives from survey respondents were selected for presentation.

3. Results

3.1. Characteristics of Study Cohorts

With the exception of a slightly higher undergraduate science grade point average (GPA) in the DO23 class, the 2018 and 2019 study cohorts showed no statistically significant differences (Table 1).

Table 1. Characteristics of study cohorts of the Des Moines University 2018 and 2019 Doctor of Osteopathic Medicine (DO) and Doctor of Podiatric Medicine (DPM) classes (averages and standard deviations), * denotes statistically significant differences between 2018 and 2019.

Class/Year	N	Overall GPA	Science GPA	MCAT	Female to Male Ratio	Age at Matriculation (Years)	Years since First Degree
DO22/2018	218	3.61 ± 0.015	3.54 ± 0.17	507.0 ± 0.27	0.44	24.6 ± 0.15	2.1 ± 0.13
DO23/2019	217	3.60 ± 0.015	3.55 ± 0.17	507.9 ± 0.27 *	0.42	24.7 ± 0.14	2.2 ± 0.12
DPM22/2018	53	3.44 ± 0.043	3.33 ± 0.048	495.3 ± 0.73	0.45	24.5 ± 0.36	1.62 ± 0.22
DPM23/2019	53	3.46 ± 0.018	3.32 ± 0.022	496.2 ± 0.33	0.32	24.9 ± 0.25	1.96 ± 0.17

3.2. Academic Performance

The introduction of B2B testing was accompanied by a significant increase in item difficulty in MCTB exams, while BMG exams were less strongly affected (482 questions analyzed, 2018 cohort size 271, 2019 cohort size 270; Figure 2). The fractions of correct responses in three MCTB exams dropped significantly from 2018 to 2019, and the average drop over all four paired exams was significant at $5.8 \pm 1.2\%$ ($p = 0.002$). This was in contrast to the BMG exam items, for which difficulty increased significantly in only one exam, and the average drop in the paired exams was not significant ($1.0 \pm 1.45\%$, $p = 0.23$). In each significant case, the increase in item difficulty was accompanied by an increase in average point biserial (as measure of discrimination) and KR20 exam reliability. The drop in examination scores was large and significant enough to trigger a score adjustment (as specified in the course syllabi). The data presented in this study are pre-adjustment; however, it cannot be excluded that students’ expectations of score adjustments influenced study patterns, and so added a confounding effect to the analysis. The effect size of the change in exam performance is demonstrated by its effect on course grades. Exam percentages are converted into letter grades by the end of the course, and as a consequence of lower exam performance, grades in the MCTB and BMG courses showed a significant drop from 2018 to 2019 (Figure 3). The median BMG grade dropped from 86% to 83% (both percentages reported as a grade of B on the transcript; average dropping from $84.8 \pm 7.8\%$

to 83.1 ± 8.4), while the median MCTB grade dropped from 90.0% to 84.0% (from reported grade of A- to B; average dropping from 88.3 ± 6.7 to $83.1 \pm 8.8\%$).

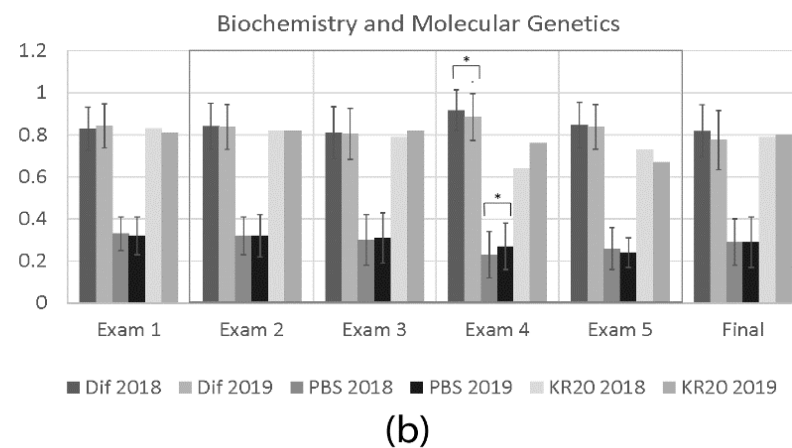
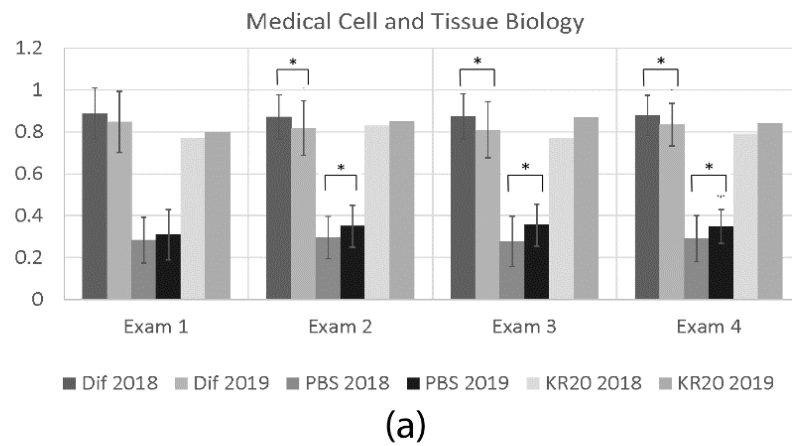


Figure 2. Performance parameters of identical exam items in Medical Cell & Tissue Biology and Biochemistry/Molecular Genetics courses in 2018 (SO) and 2019 (B2B). (a) Performance parameters for MCB exams. (b) Performance parameters for BMG exams. The box denotes exams 2–5 that were administered B2B with the MCTB exams. DO and DPM cohorts were combined. * denotes significant differences between 2018 and 2019 ($p < 0.05$). Dif, item difficulty expressed as fraction of correct answers; PBS, point biserial; KR20, exam reliability calculated as Kuder–Richardson 20 value.

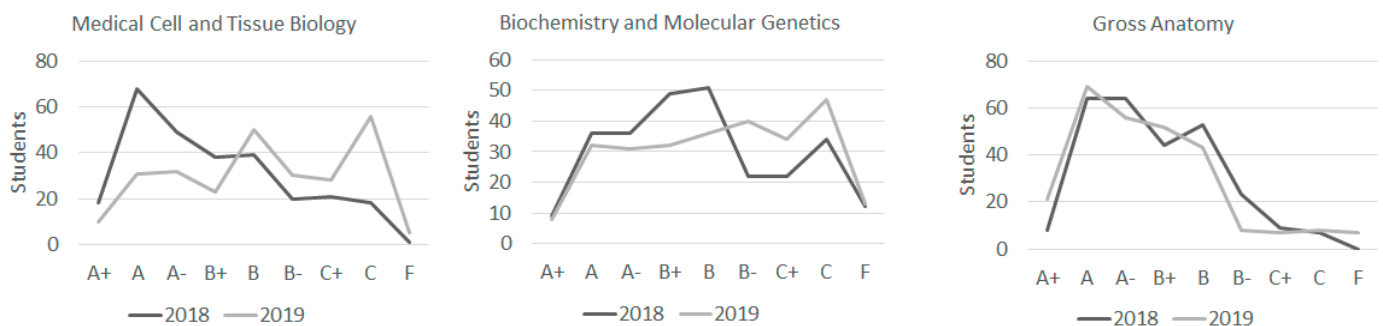


Figure 3. Grade distribution in the 2018 and 2019 offerings of Medical Cell & Tissue Biology (MCTB), Biochemistry and Molecular Genetics (BMG) and Anatomy courses (DO and DPM students, total cohort sizes 271 and 270 students, respectively). The grade distributions in MCTB and BMG courses showed a significant shift to lower averages and medians after the introduction of B2B exam scheduling. Grades in the Anatomy course (that had not undergone exam scheduling changes) showed no difference (significance determined by Mann–Whitney U testing of rank distribution).

The following controls were included in the analysis. First, to control for possible global variations in cohort performance, exam item parameters and course grades of the Gross Anatomy course were examined for significant differences between the 2018 and 2019 cohorts. Students are required to take this semester-long course in addition and parallel to the MCTB and BMG courses, and the exam structure and difficulty of the Anatomy course is similar to the MCTB/BMG offerings. No significance was found in Anatomy exam item performance, and the overall course grades were largely unchanged from 2018 to 2019 (the increase in course average by 0.1% was not statistically significant; Figure 3). Second, schedule-independent variations in the 2018/2019 cohorts' Biochemistry/Genetics performances were estimated by comparing exam item parameters of the first, unpaired BMG exams (Figure 2). The cohorts' performances on the unpaired BMG exam were not significantly different between 2018 and 2019 (even slightly better in 2019, at 84.6% vs. 83.2%). Third, to put the effect size of the scheduling-associated performance difference into perspective, we analyzed item and course grade variation between 2017 and 2018—years during which no changes were made to course content or examination schedule. No significant changes were detected in any of the parameters, most clearly illustrated by the observation that course averages fell only by 0.3% in MCTB ($p = 0.282$, compare to a scheduling-associated fall of 5.2%) and 0.5% for MG ($p = 0.091$, compared to the scheduling-associated decrease of 1.7%). Effect sizes for differences in course grades were determined using Cohen's d [20]. Effects of the 2018/2019 SO-B2B transition were found to be significant in both courses; they were categorized as small for the BMG course (0.209) and medium-sized in the case of the MCTB course (0.665). The differences in the controls (2017/2018 iterations of the courses) were not significant (0.066 for BMG and 0.098 for MCTB).

The available control data suggest that the academic abilities of the 2019 student cohorts were comparable to or even better than those of their 2018 counterparts. The controls also suggest that the difference in academic performance between 2018 and 2019 was limited to the exams for which scheduling had been changed to the B2B model, and the comparison with the respective 2017 performance data suggests that the scheduling-associated effect was larger than the normal variation between academic years.

3.3. DO Student Satisfaction with Curriculum and Testing Schedule

First-year osteopathic medical student responses to the 2019 B2B testing schedule were far more negative than the comments made on the 2018 SO exam schedule (Table 2). While members of the 2018 cohort suggested fewer, more spaced-out exams or the introduction of a true block schedule as curricular improvements, the 2019 cohort criticized the B2B testing schedule as broadly nonfunctional, poorly coordinated in terms of content, not in line with the expectations of true block testing, and stressful/exhausting. One of the purported advantages of B2B scheduling presented during orientation—the opportunity to build cognitive stamina for licensing examinations—was not mentioned by the survey respondents. Several respondents commented on the futility of their efforts to provide feedback on the B2B schedule, which can be seen as evidence that survey respondents intended to provide constructive input related to their experiences.

Table 2. Responses on DO Student Survey Reports 2018/2019. First-year DO student free-text suggestions for curricular improvements in relation to scheduling/attendance were analyzed for comments on the current testing schedule. * Cohen’s kappa for inter-rater reliability 0.84. ** Some responses were found to address more than one coded theme (in 2019, 44 respondents generated 51 comments).

	Survey of DO22 (2018, SO)	Survey of DO23 (2019, B2B)
Total respondents to survey	149 (67% of class)	135 (63%)
Respondents commenting on curriculum/scheduling (RCC)	45	44
RCC commenting on exam	44% (20/45)	95% (42/44)
RCC commenting negatively on exam *	36% (16/45)	86% (38/44)
Common negative comments by RCC, coded by theme and ranked ** (cutoff off 5%)	“Need fewer exams” (18%, 8/45) “Introduce Block Schedule” (11%, 5/45) “Exams need to be spread out” (9%, 4/45)	“B2B schedule does not work” (36%, 16/44) “Material on stacked tests is not coordinated” (20%, 9/44) “Block schedule needs to be implemented properly” (20%, 9/44) “B2B testing is stressful/exhausting” (13%, 4/44) “Respect student feedback” (9%, 4/44)

4. Discussion

The question of the optimal scheduling of exams and other cognitive tasks is of great importance for educators, as simple changes in exam schedule can influence the measured performance of examinees [2,8]. Learner performance evaluation is of particular importance in medical education, where a demanding curriculum with frequent high-stakes examinations creates a competitive environment that has long been understood to contribute to the high prevalence of anxiety, stress, depression and burnout among medical students [21,22]. Given that long, frequent and challenging examinations are a major driver of medical student stress [1], curricular reforms at medical schools encompass novel assessment strategies that aim to reduce testing stress and are designed to favor a deep understanding of biomedical concepts over the short-term memorization of facts [9].

A time-tested approach to improve the learning experience in medical education is the bundling of related assessments in a systems-integrated block schedule model [3,10]. In systems-integrated scheduling (SIS), curricular content is coordinated in a systems-based format to allow the integration of topics across related courses and the assessment of learning in fewer, integrated examinations. It was argued that the SIS of exams in a block frees time for uninterrupted study, and that the simultaneous assessment of related subjects would foster knowledge integration, as it requires students to stay current in multiple related courses—particularly when the assessment features integrated questions covering several knowledge domains [3]. However, since the examination schedule is usually only one element of a broader curriculum reform [9], it is difficult to separate the effects of scheduling from the effects of other measures to foster content integration. In addition, ethical concerns and logistical problems often preclude the execution of true randomized-control trials for the assessment of curricular interventions, limiting the availability of reliable data [23]. The present quasi-experimental study is unique and different from related studies on exam scheduling, in that it examines the effects of exam stacking in the absence of other curricular reforms, and thus provides novel insights into the scheduling effects of medical school exams. It was hypothesized that in spite of its fundamental differences, the B2B testing of related subjects would provide benefits similar to the established SIS model, by primarily improving measured learning outcomes and secondarily improving student satisfaction with curricular scheduling.

In this study, we demonstrate that student cohorts experiencing B2B scheduling of exams in Biochemistry/Molecular Genetics (BMG) and Medical Cell & Tissue Biology (MCTB)—two related courses in the first year of the medical programs at Des Moines University—show lower academic performance and lower levels of satisfaction with cur-

ricular scheduling than the cohorts tested with the traditional SO model. While these observations could be caused by any number of confounding factors that were not controlled in this natural experiment, it is tempting to speculate that the testing schedule is the major factor behind the negative outcomes of the intervention. A further complication would include both direct and indirect impacts related to the examination schedule. While it is possible that there were some benefits of these changes, these appeared to be outweighed by the negative factors in this study. This result is somewhat unexpected, as a recent related study on exam timing has shown that scheduling effects on exam performance are complex, with warm-up effects balancing fatigue and distraction counteracting the benefits of recuperation [2]. In this 2020 study, the authors demonstrate that a shortened time between exams can improve measured academic performance, particularly for high-achieving students testing in STEM subjects, but it also was evident that cognitive fatigue negatively affects performance in closely spaced analytic reasoning-intensive tasks [2,8].

The thematic coding of osteopathic medicine students' comments on the annual survey suggests an explanation for the observed negative effects of exam stacking in the study cohort. A full 40% of survey respondents specified that simply stacking exams in related courses is not conducive to learning, and that the benefits of block scheduling can only be realized if courses are fully integrated (sum of comments on "poorly coordinated subjects" and "not properly implemented block schedule"). As Harden pointed out in his groundbreaking work on the integration ladder [24], curricular integration occurs on several levels. In Harden's categorization, the DMU-COM BMG and MCTB courses only reached the basic integration levels of mutual awareness and harmonization of content; no attempts towards further content integration were made to accompany the introduction of B2B scheduling. Student comments suggested that for the block scheduling of exams to have a beneficial effect, curricular content should be fully coordinated and presented in a multidisciplinary approach; student and faculty input should be sought and incorporated to ensure the success of integration efforts. This may help explain some of the differences seen in this study as compared to more unified integration efforts presented in other studies [11]. While we did not set out to explicitly determine measures of student stress, our analysis of survey data regardless suggests that one of the reasons that the B2B schedule is perceived as detrimental to learning is that it produces stress and exhaustion, as students are simultaneously preparing for exams in two challenging courses. In this way, the B2B examination approach may have limited the utility of a student examination preparation strategy that included cramming, without offering the benefits suggested by more integrated approaches.

It is worth noting that the 2019 scheduling change was accompanied by a stronger decrease in academic performance in the MCTB assessments than in the paired BMG exams. For this observation, there are two, not mutually exclusive explanations: exam timing and course policy. First, since the MCTB exams were administered after the BMG assessments, they were more strongly affected by the cognitive fatigue of examinees. The randomization of exam order could have helped to eliminate this confounding variable; however, since the study was performed as a natural experiment, the experimental conditions were not under the control of the investigator. Second, the courses had different policies regarding the remediation of poor exam performance. In contrast to the policies in the BMG course, the MCTB course offered students immediate opportunities for grade improvement through the retesting of a failed exam. Informal communications with students showed that this policy was seen as more forgiving than the BMG course policy of grade replacement, where only one poor score could be substituted with the score of the final comprehensive examination. We conclude that, when faced with an overwhelming amount of material to prepare for two courses, students prioritize efforts in the course with higher stakes related to poor exam performance and lower chances of grade adjustments. This may have allowed students to utilize more familiar examination preparation strategies that provided additional time between the exams by postponing the true assessment of MCTB knowledge to the retest.

5. Conclusions

Our data show that the introduction of B2B scheduling for MCTB and BMG exams without the proper harmonization of content and grading policies was accompanied by significant drops in student academic performance, and generated overwhelmingly negative comments on student satisfaction surveys. The study did not attempt to ascertain the influence of the B2B schedule on the retention of knowledge, pointing to the possibility that some of the purported benefits of exam bundling were missed and suggesting that further research might be needed for a comprehensive analysis of the intervention's outcome. While the limited scope of this study precludes establishment of causality and points to the necessity of further inquiries into the topic, the outcomes of our intervention caution against such research. In response to student feedback and learning outcomes data, DMU-COM abandoned the B2B testing schedule in the following spring semester as part of the myriad measures to improve the learning environment during the COVID-19 pandemic.

Author Contributions: Conceptualization, M.S.; methodology, M.S., B.P.; validation, M.S., B.P., S.W.; formal analysis, M.S., B.P.; investigation, M.S., B.P., S.W.; resources, M.S.; data curation, M.S., S.W.; writing—original draft preparation, M.S.; writing—review and editing, M.S., B.P., S.W.; visualization, M.S.; supervision, M.S.; project administration, M.S.; funding acquisition, M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and granted exempt status by the Institutional Review Board of Des Moines University (IRB#2019-17).

Informed Consent Statement: Participant consent was waived due to the exempt status of the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: Authors declare no conflict of interest.

References

1. Bedewy, D.; Gabriel, A. Examining perceptions of academic stress and its sources among university students: The Perception of Academic Stress Scale. *Health Psychol. Open* **2015**, *2*, 2055102915596714. [[CrossRef](#)] [[PubMed](#)]
2. Goulas, S.; Megalokonomou, R. Marathon, Hurdling, or Sprint? The Effects of Exam Scheduling on Academic Performance. *BE J. Econ. Anal. Policy* **2020**, *20*, 36. [[CrossRef](#)]
3. Streips, U.N.; Virella, G.; Greenberg, R.B.; Blue, A.; Marvin, F.M.; Coleman, M.T.; Carter, M.B. Analysis of the Effects of Block Testing in the Medical Preclinical Curriculum. *J. Int. Assoc. Med. Sci. Educ.* **2006**, *16*, 10–18.
4. Preston, R.; Gratani, M.; Owens, K.; Roche, P.; Zimanyi, M.; Malau-Aduli, B. Exploring the Impact of Assessment on Medical Students' Learning. *Assess. Eval. High. Educ.* **2020**, *45*, 109–124. [[CrossRef](#)]
5. Shah, M.; Hasan, S.; Malik, S.; Sreeramareddy, C.T. Perceived stress, sources and severity of stress among medical undergraduates in a Pakistani medical school. *BMC Med. Educ.* **2010**, *10*, 2. [[CrossRef](#)]
6. Shumway, J.M.; Harden, R.M. AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Med. Teach.* **2003**, *25*, 569–584. [[CrossRef](#)]
7. Jayakumar, K.; Lapin, J.; Kogan, J. Medical Students' Preferences Regarding Preclinical Exam Scheduling: Implications for Improving Student Well-being. *MedEdPublish* **2016**, *5*. [[CrossRef](#)]
8. Pope, D.; Fillmore, I. The impact of time between cognitive tasks on performance: Evidence from advanced placement exams. *Econ. Educ. Rev.* **2015**, *48*, 30–40. [[CrossRef](#)]
9. Pock, A.R.; Durning, S.J.; Gilliland, W.R.; Pangaro, L.N. Post-Carnegie II curricular reform: A north American survey of emerging trends & challenges. *BMC Med. Educ.* **2019**, *19*, 260. [[CrossRef](#)]
10. Muller, J.H.; Jain, S.; Loeser, H.; Irby, D.M. Lessons learned about integrating a medical school curriculum: Perceptions of students, faculty and curriculum leaders. *Med. Educ.* **2008**, *42*, 778–785. [[CrossRef](#)]
11. Hudson, J.N.; Tonkin, A.L. Evaluating the impact of moving from discipline-based to integrated assessment. *Med. Educ.* **2004**, *38*, 832–843. [[CrossRef](#)] [[PubMed](#)]
12. Thompson, A.R.; Braun, M.W.; O'Loughlin, V.D. A comparison of student performance on discipline-specific versus integrated exams in a medical school course. *Adv. Physiol. Educ.* **2013**, *37*, 370–376. [[CrossRef](#)]

13. Brauer, D.G.; Ferguson, K.J. The integrated curriculum in medical education: AMEE Guide No. 96. *Med. Teach.* **2015**, *37*, 312–322. [[CrossRef](#)] [[PubMed](#)]
14. El Ansari, W. Factors associated with students' satisfaction with their educational experiences, and their module grades: Survey findings from the United Kingdom. *Educ. Res.* **2011**, *2*, 1637–1647.
15. Grayson, J.P. The Relationship between Grades and Academic Program Satisfaction over Four Years of Study. *Can. J. High. Educ.* **2004**, *34*, 1–34. [[CrossRef](#)]
16. Ackerman, P.L.; Kanfer, R. Test length and cognitive fatigue: An empirical examination of effects on performance and test-taker reactions. *J. Exp. Psychol. Appl.* **2009**, *15*, 163–181. [[CrossRef](#)]
17. Hashmat, S.; Hashmat, M.; Amanullah, F.; Aziz, S. Factors causing exam anxiety in medical students. *J. Pak. Med. Assoc.* **2008**, *58*, 167–170.
18. American Association of Colleges of Osteopathic Medicine. *Osteopathic Core Competencies for Medical Students*; American Association of Colleges of Osteopathic Medicine: Chevy Chase, MD, USA, 2012.
19. McHugh, M.L. Interrater reliability: The kappa statistic. *Biochem. Med.* **2012**, *22*, 276–282. [[CrossRef](#)]
20. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*; Routledge Academic: New York, NY, USA, 1988.
21. Ludwig, A.B.; Burton, W.; Weingarten, J.; Milan, F.; Myers, D.C.; Kligler, B. Depression and stress amongst undergraduate medical students. *BMC Med. Educ.* **2015**, *15*, 141. [[CrossRef](#)]
22. Dyrbye, L.N.; Thomas, M.R.; Shanafelt, T.D. Systematic Review of Depression, Anxiety, and Other Indicators of Psychological Distress among U.S. and Canadian Medical Students. *Acad. Med.* **2006**, *81*, 354–373. [[CrossRef](#)]
23. Gopalan, M.; Rosinger, K.; Ahn, J.B. Use of Quasi-Experimental Research Designs in Education Research: Growth, Promise, and Challenges. *Rev. Res. Educ.* **2020**, *44*, 218–243. [[CrossRef](#)]
24. Harden, R.M. The integration ladder: A tool for curriculum planning and evaluation. *Med. Educ.* **2000**, *34*, 551–557. [[CrossRef](#)] [[PubMed](#)]