Competing Hypotheses About Online Versus Face-To-Face Course Outcomes: Analysis Leveraging the Unique Circumstances of the COVID-19 Pandemic

Claire Wladis
Borough of Manhattan Community College and the Graduate Center at the City University of New York

Alyse C. Hachey
The University of Texas at El Paso

Catherine A. Manly
Fairleigh Dickinson University

Katherine M. Conway
Borough of Manhattan Community College at the City University of New York

Abstract

Due to ethical and practical concerns, no large randomized controlled trials comparing outcomes in fully online versus face-to-face courses have occurred. However, circumstances presented by the COVID-19 pandemic created a unique opportunity to explore hypotheses about the relationship between online medium and course outcomes. The City University of New York (CUNY) was one of the first U.S. universities to shift courses completely online. This study used a dataset of all courses taken by students enrolled in fall 2019 or spring 2020 at any of the two or four-year colleges of CUNY to explore the relationship between course medium and completion rates. Findings from a quasi-experimental fixed effects approach show that at both two- and four-year colleges, students were significantly less likely to successfully complete fully online than in-person courses; however, this appeared to be driven more by the student's original choice of medium than the medium in which the course was actually taught. Results did not support hypotheses that the online medium itself causes lower completion; within the same term, students were less likely to successfully complete courses that they originally chose to take fully online than in-person, regardless of whether those courses were subsequently taught in-person or fully online after pandemic onset. Students also took different types of courses in different mediums, which made comparing outcomes across mediums problematic; results show students were significantly more likely to take elective courses fully online and less likely to take major requirements fully online. Findings indicate that elective courses had significantly lower rates of successful completion than major requirements, which may account for some of the differences in outcome by chosen (rather than enacted) medium.

Keywords: online versus traditional course completion, COVID-19 pandemic, elective versus major requirement course completion, online outcomes, fixed effects modeling

Wladis, C., Hachey, A.C., Manly, C.A. & Conway, K.M. (2025). Competing hypotheses about online versus face-to-face course outcomes: Analysis leveraging the unique circumstances of the COVID-19 pandemic. *Online Learning, 29*(3), 176-210. https://doi.org.10.24059/olj.v29i3.5018

Introduction

A correlation between online course medium and negative outcomes is often implicitly assumed in higher education discourse, with online viewed as inferior to face-to-face (Jaggars & Bailey, 2010; Protopsaltis & Baum, 2019). Some contend that the online medium itself causes worse outcomes (Altingdag et al., 2021; Baum & McPherson, 2019), while others argue that students who choose to take courses online take different courses (Bejerano, 2008; Fischer et al., 2021; Jaggars, 2014; Wladis et al., 2014, 2017) or have different characteristics than those who do not (Fischer et al., 2020; Hachey et al., 2022; Hyllegard et al., 2008; McDonough et al., 2014; Wladis et al., 2018, 2022). New knowledge is needed to shed light on the accuracy of each of these three hypotheses, as the connotations of each have important higher education policy implications for the efficacy of online learning in the post-COVID-19 era. However, large, representative randomized controlled trials comparing outcomes in fully online versus face-to-face courses are not feasible due to ethical and practical concerns.

With no large, randomized evidence available, the circumstances of the COVID-19 pandemic created a unique opportunity to explore hypotheses about the relationship between the online course medium and course outcomes. The City University of New York (CUNY) was one of the first U.S. public universities to shift all courses completely online due to the pandemic. This allowed for a within-student comparison of outcomes in courses that were forced to transition online to those courses that were originally online at the beginning of the spring 2020 semester. We know of no other large-scale *within-subjects* study that controls for the impact of self-selection into online courses by separating the *originally chosen mode* of the course from the *actual mode* of the course as subsequently implemented. To address this research gap, we compared such courses *within the same student* taking multiple courses originally chosen in different mediums *during the same term*, making this a novel study design.

Given the evolving higher educational landscape, where current data indicates that post-COVID-19 student preferences have shifted from a clear preference for in-person classes to a more equal distribution across modalities and wider exposure/experience of faculty to online learning (Bright & Vogler, 2024), a better understanding of the presumed relationships between the online course medium and course outcomes is critically needed for institutional decision-making and optimal utilization (Alarifi & Song, 2024). This study helps provide missing knowledge by leveraging the unique circumstances of the shift to fully online instruction during the COVID-19 pandemic, shedding new light on the patterns between online course taking and course outcomes. While this research is observational and not causal, it allows for exploration of the extent to which observed patterns are consistent with existing hypotheses about the online medium using a quasi-experimental fixed effects approach that is a stronger design than typical regression-based studies.

Literature Review

To put this study and its hypotheses in context, existing literature offers background for understanding patterns connecting completion outcomes to online course-taking, course characteristics, and student characteristics. This study draws from a long history of research on outcomes in online courses. First, we consider the research on online course outcomes, outside of the context of emergency remote teaching (ERT). Then, we discuss how student choices to take different types of courses online versus face-to-face may explain some of the patterns observed with course outcomes by medium. Next, we describe how students' characteristics may impact their choice of course medium, which may also explain some patterns observed in course outcome by medium. Finally, given the context of this study, we review research on outcomes during ERT, such as during the recent COVID-19 pandemic. All of this sets the stage for the subsequent conceptual framework that orients our study analysis.

Outcomes in Online Courses

Pre-pandemic, some studies indicate better retention/graduation rates for students who take online courses or no-difference between instruction mediums (Shea & Bidjerano, 2014; James et al., 2016; Johnson & Mejia, 2014). Others found lower retention/graduation rates for students taking online courses versus only face-to-face courses (Jaggars & Xu, 2010; Smith, 2016; Xu & Jaggars, 2014). Results are similarly mixed for analysis of course grades, with some studies reporting higher grades online, some no difference in grades between medium, and some worse grades online (for a review, see Hachey et al., 2022). It has been suggested that while proximal outcomes in online learning may be negative compared to face-to-face learning (i.e., course completion; course grades), benefit can be found in online course-taking more distally (i.e., time-to-degree; degree completion) (Castro & Tumibay, 2021; Fischer et al., 2021). Yet, there is a call for a more nuanced analysis of course outcomes, as the mixed findings in the literature are likely a result of significant contextual and methodological differences found between past empirical studies (Alarifi & Song, 2024; Hachey et al., 2022).

Somewhat consistent in the literature is that students are more likely to drop online than face-to-face (Hachey et al., 2022). This is problematic, as evidence strongly suggests that course completion impacts academic momentum, which is tied to persistence and degree completion (e.g., Chan & Wang, 2018). These findings have led to a perception that the online medium itself negatively impacts students (Altingdag et al., 2021; Baum & McPherson, 2019; Protopsaltis & Baum, 2019). A study comparing pre-pandemic dropout reasons found that students in both mediums were equally likely to cite course characteristics as a reason for dropout, but students in online courses were also more likely to cite reasons related to lack of time (Hachey et al., 2023). Research suggests that a lack of *time capital* (i.e., time as a resource for college) is a factor in online enrollment and outcomes (Hachey et al., 2023; Wladis et al., 2022). Students pressed for time may be more likely to drop because they are also more likely to take courses that are less important to their degree online than face-to-face (Jaggars, 2014; Wladis et al., 2014, 2017). This may be particularly salient for women and students of color who have on average lower time capital available for college (Wladis et al., 2022, 2024).

Course Characteristics and Online Outcomes

Students tend to choose to take different types of courses online than face-to-face. Bejerano (2008) and Jaggars (2014) report that students chose to enroll in more "interesting" or "important" courses face-to-face rather than online. Evidence also suggests that students prefer to take different subjects face-to-face versus online (Jaggars, 2014; Mann & Henneberry, 2014). Because outcomes may be different in different types of courses, correlation between online medium and course type could be confounding differences between online and face-to-face course outcomes in the literature. For example, online engagement has been found to be different in online courses across different subject areas (Young & Bruce, 2011) and research both during (e.g., Tsang et al., 2021) and before (e.g., Jaggars & Xu, 2014) the pandemic indicated that engagement can impact perceived and actual outcomes in online courses. In fact, Xu & Jaggars (2014) found variation in online versus face-to-face performance across subject areas, indicative of differences in the types of courses students chose to take online versus face-to-face driving some differences in online versus face-to-face course outcomes. Further, Wladis et al. (2014, 2017) found that electives in comparison to major requirements were more likely to have a larger gap in successful course completion rates online versus face-to-face. Taken together, the research suggests that the fact that students take different types of courses online than face-to-face may also impact their choices about which courses to prioritize academically. For example, when students find themselves pressed for time (a particularly common scenario given the higher prevalence of time poverty among online students [Wladis et al., 2022]), they may be more likely to invest less time (thus risking failure) in those courses or drop those courses that are least "interesting" or "important" to them, which some research suggests are more likely to be the courses that they chose to take online.

Student Characteristics and Online Outcomes

Differences in outcomes between online and face-to-face sections may also be the result of the characteristics of those who enroll in each medium (Hyllegard et al., 2008; McDonough et al., 2014). During typical enrollment (i.e., outside of pandemic conditions), students choose their course modality, such that students who take online courses have selected the online medium, and the specific courses to be taken in that medium online versus face-to-face, voluntarily. Students who enroll online often need flexibility or have less time to devote to college than those who do not (for example, because they have work and/or family responsibilities), and available time capital has been found to explain differences in outcomes by course medium (Daymont et al., 2011; Jaggars et al., 2013; Wladis et al., 2022). Students who enroll online have different demographic characteristics than those who enroll face-to-face; online students are more likely to be older, to be women, and to have non-traditional characteristics such as working full-time and single parent status (Johnson & Mejia, 2014; Pao, 2016; Xu & Jaggars, 2013; Wladis et al., 2018, 2022). Further, differences in academic preparation including remedial placement and GPA, race/ethnicity, gender and age have been linked to online course outcomes (see review in Hachey et al., 2022). Yet, much of the existing literature does not control for student characteristics (Smith, 2016), and those that do tend not to account for differences in course characteristics. Because student-level characteristics have been found to impact results (e.g., Alarifi & Song, 2024; Fischer et al., 2021), it is difficult to interpret the results of previous research that did not account for student characteristics; the present study specifically addresses this by employing a within-student design.

Outcomes in Online Courses in Emergency Remote Teaching (ERT)

ERT is online learning adopted in response to a disaster or crisis to maintain educational continuity; in this context, students find themselves in a modality that they did not select (Gelles et al., 2020; Hodges et al., 2020). While extensive research related to ERT during the COVID-19 pandemic has occurred, in a metareview of this type of research, Doo et al. (2023) report that only 4% of studies across 195 countries focused on student outcomes. Therefore, this study adds to the very little research that specifically investigates student outcomes during the COVID-19 pandemic.

The studies that have analyzed student outcomes during ERT have focused on the overall effects of ERT itself, rather than exploring how the unique circumstances of ERT can be leveraged to explore how course medium may relate to outcomes for those students who elect to take online courses during typical circumstances. McDougall et al. (2020) found that students at four research intensive institutions had significantly less achievement on standardized assessments during ERT versus pre-pandemic. Bird et al. (2020) reported a 6.7 percentage point decrease in course completion for two-year students in the Virginia Community College system who started out face-to-face and shifted online during the pandemic. Kofoed et al., (2021) found that West Point students randomly assigned to online courses during ERT had lower grades than those randomly assigned to face-to-face sections. However, in contrast, Altingdag et al. (2021) found that both online and face-to-face students had better course outcomes during spring 2020 compared to spring and fall 2019 (such findings are often attributed to institutional grading policy changes during the COVID-19 pandemic [see Chan, 2023]). What all of these studies have in common is that they focus on the impacts of ERT across the entire student population, including students who did and did not chose online courses voluntarily. That focus allows us to use these studies to make inferences about when and whether students might be harmed by ERT, or by being forced into online courses that they did not choose.

However, these studies are inappropriate for drawing inferences about students who *elect* to take courses online voluntarily, or for guiding policies and structures related to online course-taking outside of ERT. In prior work, we found that student course outcomes during ERT were highly dependent upon whether they had elected to take online courses before the pandemic started. For example, we (Wladis et al., 2023) found that students at two-year colleges, men, and Black and Hispanic students had worse outcomes during ERT, but only if they had *not* chosen to enroll in online courses pre-pandemic. In this study, a similar approach is taken: we use the unique circumstances of ERT during the pandemic *not* to compare outcomes in online courses for *all* students, but rather to compare course outcomes by delivered medium just for the subset of students who ordinarily *voluntarily* chose online courses. This means that outcomes from this study can inform policies most relevant to those students who actually chose to take online courses in college.

Different Target Populations for Online Outcome Research

Research on course outcomes during ERT points to a critical question about the target population when pursuing research on online course outcomes. Many studies leave the target population underspecified, but existing research on online course outcomes tends to point implicitly to one of two categories of target population: (a) the entire student population (i.e.,

where average effects across all students are measured under conditions where students are assigned to online courses involuntarily [e.g., during ERT or in Randomized Controlled Trial {RCT} study designs]); and (b) the population of students who chose to enroll in online courses (i.e., designs where only online course outcomes for students who selected them voluntarily are studied). Research results using different target populations provide very different inferences, leading to different practical implications from research outcomes. For example, a few studies have performed random assignment of students in online versus in-person courses (pre-pandemic: Alpert et al., 2016; Arias et al., 2018; Figlio et al., 2013; Stack, 2015; during pandemic: Kofoed et al., 2021), with mixed results. These studies are not sufficient for making inferences about whether online learning is good/bad for the subset of students who *choose* to take courses online, since this population has different characteristics than the general population, and student characteristics have been shown to matter, as online learning may be beneficial to some students yet harmful to others (Alarifi & Song, 2024; Wladis et al., 2022).

Course medium selection is not random; under normal circumstances, a subset of students *elect* to take courses online for compelling reasons, most often to address life or time factors that make attending college in-person difficult (i.e., scheduling conflicts with work and/or childcare responsibilities; long commutes; physical disability); the issues that often drive students to self-select into online courses may also make them more likely to dropout due to other life commitments taking resources away from their college studies (Bejerano, 2008, Jaggars, 2014; Wladis et al., 2014, 2015, 2022, 2023). A randomized-assignment study with this population would select students already enrolling in an online course and randomize the specific course section taken online versus face-to-face; this would eliminate course-characteristics-selection effects. However, this approach is not practical and as such, we could not find any such studies in the research literature. In the absence of randomized-assignment, one option (which we utilize here) is to use observational data from the switch to ERT during the COVID-19 pandemic to compare the within-student outcomes of courses that were originally online to those that were originally face-to-face, for students who had chosen to enroll in online courses prior to the start of the pandemic. Because of the within student comparison, this study design utilizes a quasi-experimental fixed effects approach and naturally controls for student-level, course-level and self-selection confounds; thus, it serves to provide a novel empirical contribution.

Conceptual Framework

Existing research is often used to drive higher education discourse. As an impetus for informing policies around online course-taking, the literature points to multiple hypotheses about factors that impact student outcomes in online versus face-to-face courses, each with different implications for practice:

Competing Hypotheses: *When observed, lower rates of successful course completion in online versus face-to-face courses are caused by:*

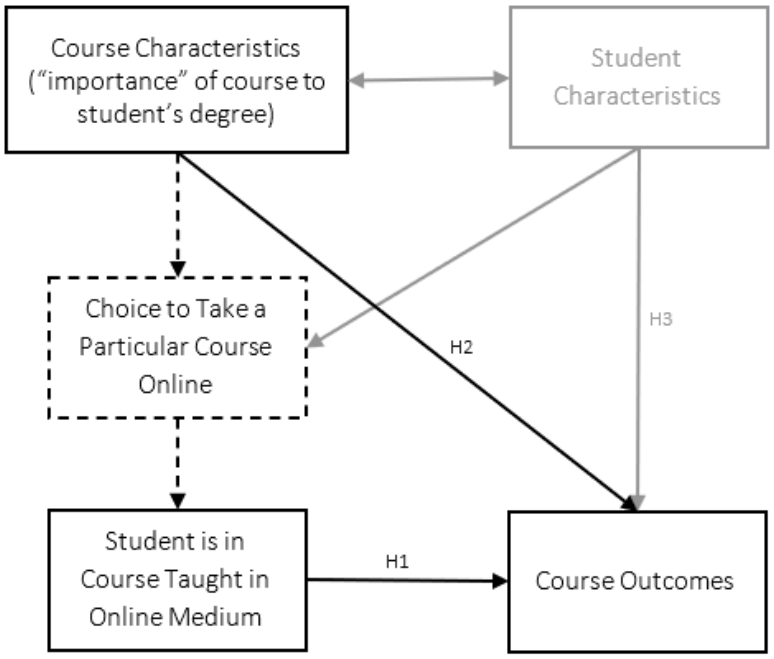
- ***H1.*** *The online medium itself (e.g., Altingdag et al., 2021; Baum & McPherson, 2019).*
- ***H2.*** *Differences in the types of courses that students choose to take online versus those that they take face-to-face (e.g., Bejerano, 2008; Jaggars, 2014; Wladis et al., 2014).*
- ***H3.*** *Differences in characteristics between students who choose to take courses online versus those who do not (e.g., Hyllegard et al., 2008; McDonough et al., 2014;*

Wladis et al., 2022).

Our population of interest is explicitly limited to the subset of students who choose to take courses online, so H3 is not relevant in this study (students who did not choose to take at least one online course were excluded from the sample). This is not to contend that H3 is not a real factor impacting online course outcomes, but rather to acknowledge its importance; indeed, mounting evidence both pre- and post-pandemic suggest that student characteristics and their life contexts likely impact both enrollment and dropout online (Alarifi & Song, 2024; Fischer et al., 2021; Hyllegard et al., 2008; Johnson & Mejia, 2014; McDonough et al., 2014; Pao, 2016; Xu & Jaggars, 2013; Wladis et al., 2018, 2022). However, what is missing in the literature (and thus the aim of this study) is an exploration of H1 and H2 while fully controlling for student characteristics and selection bias (H3). For this reason, we compare outcomes between courses taken by the same student in the same term, so the models explored here remove potential variation caused by H3. H1 and H2 may be possible mechanisms explaining the outcomes of students who elect to enroll in online courses outside of pandemic conditions, so in this study we leverage unique circumstances of ERT during the COVID-19 pandemic to explore the extent to which patterns in the data are consistent with each of these hypotheses. We take advantage of the fact that during ERT, the chosen medium and actual implemented medium of courses differed, allowing us to separate out potential impacts of how students choose online courses (and the subsequent impact that may have on outcomes) from potential impacts of the online medium itself. Looking at H1 and H2 together should aid in disentangling these factors' influence. This could help explain conflicting conclusions found in prior research (for a review, see Hachey et al., 2022). Figure 1 illustrates our conceptual model.

Figure 1

Model of How Course Characteristics, Student Characteristics, and the Online Medium Itself May Contribute to Course Outcomes



Notes. Grey lines represent characteristics that are controlled through fixed effects (comparing only courses within students in the same term) and by focusing only on outcomes for students who chose to enroll in online courses prior to pandemic onset.
Dotted lines represent the path by which course characteristics may influence student decisions to take a course online, which became separate from the medium in which the course was taught when all courses were moved fully online during ERT.
Path "H1" is the path denoting hypothesis 1 (the online medium itself lowers course outcomes).
Path "H2" is the path denoting hypothesis 2 (poorer outcomes in online courses are significantly related to the student's reasons for taking a course online—i.e., the course is less important to them).
Path "H3" is the path denoting hypothesis 3 (poorer outcomes in online courses are significantly related to the characteristics of the students who choose to take courses online; this hypothesis is controlled for in this study because all comparisons are within-students).

Considering these competing hypotheses together in a single conceptual model has not been previously described in the literature. This is largely because removing the effects of course medium choice is impractical and ethically problematic outside of ERT conditions. Students who take courses both online and face-to-face have been found to choose to take courses that are less "interesting" or "important" to them online (Bejerano, 2008; Jaggars, 2014; Mann & Henneberry, 2014); these courses are therefore already likely to have worse outcomes regardless of medium. Thus, outside of ERT, where students choose course medium themselves, it is impossible to disentangle the relationship between course characteristics and course outcomes from the relationship between course characteristics and course medium choice. In other words, during typical circumstances, the relationship between course medium and course outcomes is a combination of H1 pathway and the combined effects of the dotted and H2 pathways in Figure 1, and these two pathways (H1 and H2) cannot be teased apart.

However, the unique circumstances of the COVID-19 pandemic removed the relationship between chosen medium and actual medium, as all courses were moved fully online (regardless of the medium that the student had originally chosen for them). Pandemic conditions thus allowed us to disentangle the two paths H1 and H2. Comparing outcomes in courses that a student originally chose in the online versus face-to-face medium allowed us to explore H2 separately (since both of these types of courses were actually delivered in the same medium during the pandemic); hence, any remaining differences in outcomes between these two types of courses (taken by the same student in the same term) can be isolated to those course characteristics that drove the student to choose the online versus face-to-face mediums for particular courses over others. Because all courses were delivered in the same medium, differences in outcomes in this situation cannot be attributed to the online medium itself (H1).

We note that it is possible that both H1 and H2 play a role—we do not contend that these two hypotheses are necessarily mutually exclusive; existing research does not definitively suggest one may have precedence over the other or that they only may have individual impact. Here, we investigate the extent to which data obtained during the unique circumstances of the COVID-19 pandemic exhibits patterns that are or are not consistent with each hypothesis (H1 and H2), allowing for interpretation of our study results in light of each of these hypotheses. This conceptual framing enables us to begin to deconstruct the simplistic framing of some existing research that has considered course medium without other potentially relevant factors. Interpreting our results through this more nuanced lens serves to offer scholarly and practical

insights that go beyond existing literature. The present study uses this conceptual model to frame several hypotheses for empirical investigation of within-student data through quasi-experimental methods designed to extend our understanding beyond what previous studies have achieved through cross-sectional or between-student designs. It also leverages pandemic data in a novel way to disentangle the relationship between course characteristics, course medium choice, and course outcomes, something that has not yet been attempted in existing research literature using a within-students design.

Method

Sample and Data Source

The dataset consisted of all courses taken by students enrolled in either fall 2019 or spring 2020 at CUNY, the largest urban university system in the U.S. (Boland, 2021).[1] The study sample is limited to students who were enrolled in at least one online and at least one traditional (face-to-face) course in fall 2019 and in spring 2020 ($n = 83{,}049$), allowing for comparison across medium and term within students. By analyzing fixed effects within students, we are able to assume characteristics of the individual, which might otherwise confound the results, are the same across comparisons made with the same individual. This allows us to make comparisons across medium and across term, all for the same individual, thus offering a stronger study design than the more typical design that makes comparisons across different individuals. This is particularly critical when comparing course outcomes by medium, since students who enroll in online courses have been found to differ significantly from students who take only face-to-face courses (e.g., Alarifi & Song, 2024; Fischer et al. 2021).

In fall 2019, CUNY undergraduates were 32% Hispanic; 26% Black, 22% Asian and Pacific Islander, and 20% White. About 44% were first-generation college students, 39% were non-native English speakers, 25% were older (>25 years) and 60% were Pell recipients. More than 60% of students worked and 33% attended part-time. While CUNY is not nationally-representative because of its large diverse student body and mix of two- and four-year colleges, it is a logical choice for exploring the relationship between online learning and outcomes for a diverse student population. When CUNY announced in March 2020 that all on-campus instruction was suspended, the majority of students had not enrolled in fully online courses, and most faculty had never taught online. During the forced pandemic transition, instructors were offered extra training and technical assistance, yet they had little time to prepare and varied levels of expertise in online pedagogy.[2]

[1] One of the CUNY community colleges recoded all classes that moved online in March 2020 as "fully online" retroactively. This made these classes indistinguishable in the data from classes that were online from the beginning of the semester. Therefore, we excluded data from that college from analysis.

[2] We note that the limited preparation of faculty during ERT may have influenced course outcomes. To address this, further investigation was conducted to investigate this potential confounding relationship (see subsequent section on Exploring Potential Confounding Relationships).

Measures

Online medium was operationalized as the format that was used to deliver instruction; this is the way that the online medium is operationalized in most studies that explore its relationship to academic outcomes (e.g., Shea & Bidjerano, 2014; Johnson & Mejia, 2014). The term ***online medium*** describes courses that are delivered 100% online; it does not indicate *quality* of instruction. (We discuss how varied quality of online instruction during ERT and other pandemic conditions may play a role in observed patterns in later sections.)

Outcomes were compared in courses based on the ***chosen medium***—the medium *originally selected by the student* (prior to the pandemic). In fall 2019, the chosen medium and the actual medium were the same for all courses; in spring 2020, the medium was online for most of the semester, often departing from the student's chosen medium. We call courses that students originally selected in the online medium ***fully online courses***, and those that they did *not* originally select in the online medium ***traditional courses***. Thus, traditional courses were not fully online in fall 2019 but they were fully online for most of the spring 2020 term.[3] We compared how gaps in course outcomes between fully online and traditional courses shifted from the fall 2019 ***pre-pandemic term*** to the spring 2020 ***pandemic term***. All face-to-face instruction was suspended at CUNY on March 11, 2020, with ERT enacted. In the spring 2020 term, students may have been subject to a number of direct and indirect effects from the pandemic. However, we compared only courses within students in the same term; thus, term is a proxy for student-level pandemic effects; in this way, it serves as a control for any potential impact of student-level pandemic effects on outcomes.

The outcome of interest was ***successful course completion***—completion with a C- or better[4]; this is typically necessary for transfer credit or credit in a major. Only students who were enrolled in at least one traditional and at least one fully online course in both terms were included; this allowed comparison of course outcomes across mediums within students within terms. Control variables included: gender, race/ethnicity, age, G.P.A., first-semester freshman status, the median household income of a student's zip code, and college level (two-year versus four-year). Some CUNY colleges were on a separate calendar, so in some cases we controlled for the semester start date.

The only variable with missing values was the median household income of a student's zip code (0.06% missing); students with missing data included data entry errors but were primarily a small subset of foreign students who did not have a local address on file. Because this population likely differs from those students with a zip code on file, the missingness was missing not at random (MNAR), which means that imputation would not be appropriate. Additionally, the small proportion of missing values in only a single variable meant that substantial bias was unlikely. Thus, listwise deletion was used, because it is a reasonable choice for handling missing

[3] We compare originally chosen medium (rather than actual medium) because if students selected different types of courses online than face-to-face (e.g., by taking courses that are less important to them online as suggested by Bejerano [2008] and Jaggars [2014]), and this is correlated with differential course outcomes (H2), then this categorization allows us to disentangle H2 from H1.
[4] We include pass/fail designations in which a student was categorized as passing the course and receiving credit.

data under such circumstances, and other approaches such as imputation are only appropriate when data are reasonably assumed to be missing at random (MAR) or missing completely at random (MCAR) (Allison, 2002; Manly & Wells, 2015).[5]

For one sub-analysis to assess H2, we considered the relevance of course type to a student's major. This required hand-coding all courses taken during the 2019-2020 school year for each student in each major based on course type—whether the course counted as an elective, distributional, or major requirement. Major requirements were required courses in the major discipline or a closely related discipline (e.g., math and computer science were considered closely related)[6]; distributional requirements were courses that were required for the major, but not directly related to the major discipline (e.g., English requirements for the math major); and electives were courses that were not required by the major but fulfilled degree credit. Comparing outcomes across course type allowed us to observe patterns associated with H2.

Exploring Competing Hypotheses

We explored the following two hypotheses to determine the extent to which patterns observed in the data are consistent with H1 and/or H2 as explanatory mechanisms for differential outcomes observed in online versus face-to-face courses.

Hypothesis: *Lower rates of successful course completion in online courses are caused by:*	**Data patterns that would be consistent with each hypothesis:**
H1. *The online medium itself.*	• **Overall course outcome gap pattern expected:** Gap shrinks/closes during pandemic term (when ERT enacted). (If the online medium itself lowers successful course completion, when all courses move online, we would expect the gap in outcomes between online and face-to-face courses to disappear.) • **Robustness check:** Gap shrinks more at colleges where a greater proportion of the term occurred online. (If the online medium itself lowers successful course completion, we would expect course outcomes to be worse at colleges where there was less in-person instruction prior to the shift online; we would expect a significant interaction between time spent in in-person instruction before ERT and course medium in predicting course outcomes.)

[5] This limits generalizability of findings so that they may not apply to all first-semester foreign students, although the majority of students were still in the sample because they had a local address on file.

[6] The coding required interpretation of what disciplines were considered closely related-- faculty in those disciplines were consulted.

H2. *Students choosing to take "less important" courses in the online medium.*	• **Overall course outcome gap pattern expected:** Gap stays the same during the pandemic term (because course medium was selected prior to pandemic onset both terms). (If lower outcomes in online courses are due to differing characteristics of the courses students choose to take online versus face-to-face, then gaps in outcomes for these courses should stay the same, regardless of the actual medium in which they are delivered.) • **Robustness check:** Gap in both terms can be explained statistically by some measure of the "importance" of the course to the student's degree. (If characteristics like the "importance" of a course to a student's degree impacts both their likelihood of taking it online and the likelihood of poorer course outcomes, then mediation models should illustrate that a significant percentage of the gap in outcomes by medium is explained by measures of "importance" of the course to a student's degree plan.)

The "overall course outcome gap patterns" are intended to describe the expected trend if the corresponding hypotheses were the only factor at play. The robustness check for H1 would indicate support for H1 if there is a significant interaction such that colleges where more time was spent online saw more similar outcomes (i.e., a smaller gap) than other colleges, thereby supporting the idea instructional time in the online (rather than face-to-face) medium itself leads to lower course outcomes. The robustness check for H2 would indicate support for H2 if the "importance" of the course for the student's degree interacts significantly with course medium such that the gap was smaller for courses considered more "important," thereby supporting the idea that choosing lower "importance" courses online would lead to lower outcomes in those courses.

In subsequent analyses, we consider potential confounders and alternative hypotheses that may explain patterns observed in the data. One such confounder is ERT itself. Fully online courses and ERT are delivered similarly at a distance; however, ERT faculty have little time for planning or implementation and both faculty and students find themselves in a modality that they did not select (Gelles et al., 2020; Hodges et al., 2020). In contrast, non-ERT online courses are pre-planned and often follow research-based standards of instruction (Kebritchi et al., 2017). Thus, we might expect to see poorer outcomes in traditional courses during the pandemic term because of ERT, and not because of the online medium itself. This, and other potential confounding factors, are discussed later.

Analytical Approaches and Data Analyses

Linear probability multi-level models were used for all analyses reported here. Given the binary nature of the outcome, logistic regression models (using a logit link function) were also checked and results were compared for all analyses. Statistical analyses used Stata 17: *mixed* for multi-level logistic and linear probability models, and *melogit* for multi-level logistic regression

models. As results of both approaches were similar, only linear probability models are presented here because: (a) our primary concern was estimating particular parameters and not forecasting a specific outcome; (b) due to rescaling, we could not compare coefficients across logistic regression models (Buis, 2010), however, this could be done with linear regression; and (c) the literature illustrates how commonly odds ratios are misinterpreted (Norton & Dowd, 2018) and readers are generally apt to better interpret the coefficients of linear probability models.

The multi-level models used include three levels: (a) the individual course record; (b) the individual student; and (c) the college at which that student took that course. This allowed for controlling of clustering by student and by college, to account for the tendency of grades in different classes for the same student, and grades given for different classes at the same college, to be similar to one another.

Using *i* to index course records and *j* to index students and *k* to index colleges, our multi-level model is:

Level 1:

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}T_{1ijk} + \cdots + \beta_{njk}T_{nijk} + \varepsilon_{ijk} \quad (1)$$

Level 2. For each $l = 0, \ldots, n$:

$$\beta_{ljk} = \tau_{l0k} + \tau_{l1k}C_{l1k} + \cdots + \tau_{lmk}C_{lmk} + u_{ljk} \quad (2)$$

Level 3. For each $l = 0, \ldots, n$ and $p = 0, \ldots, m$:

$$\tau_{lpk} = \mu_{lp0} + \mu_{lp1}L_{lp1} + \cdots + \mu_{lpq}L_{lpq} + r_{lpk} \quad (3)$$

Where β_{0jk} are the level-one intercepts, $\beta_{1jk}, \ldots, \beta_{njk}$ are the level-one slopes, and ε_{ijk} represents the random effect for course record *i*, student *j*, and college *k*; τ_{l0k}; for $l = 0, \ldots, n$, τ_{l0k} are the level-two intercepts, $\tau_{l1k}, \ldots, \tau_{lmk}$ are the level-two slopes, and u_{ljk} are the random effects for student *j* at college k; and for $l = 0, \ldots, n$ and $p = 0, \ldots, m$, μ_{lp0} are the level-three intercepts, $\mu_{lp1}, \ldots, \mu_{lpq}$ are the level-three slopes, and r_{lpk} are the random effects for college k.

For logit models, the same equations were used as in linear regression, but with a logit link employed to model the probability distribution. In other words, Y_{ijk} from the Level 1 equation is replaced with $\lambda(Y_{ijk})$ where $\lambda(Y_{ijk})$ represents the logit link:

$$\lambda(Y_{ijk}) = \frac{e^{Y_{ijk}}}{1 + e^{Y_{ijk}}}$$

Here $\lambda(Y_{ijk})$ represents the probability value of the dependent variable, and other variables are the same as in the linear regression models.

Separate models were run for students enrolled at two-year versus four-year colleges; there is research establishing that student outcomes at these different types of institutions are best modeled separately (e.g., Attewell, Heil, & Reisel, 2011).

For the robustness check for H2, KHB decomposition (Kohler et al., 2011) was used to calculate direct and indirect effects for mediation analysis. This model is preferable to others because it generates an indirect effect that is not distorted by the rescaling that occurs when a potential mediator variable that is correlated with the dependent variable is added to a nonlinear model. Monte Carlo studies have shown that the KHB method always performs as well or better than other methods that have been developed to address the problems of rescaling in terms of

recovering the degree of mediation net of the impact of rescaling (Kohler et al., 2011). We recognize that the language of mediation may be associated by some with causation, and that this is an observational and not a causal study. However, attempts to reword seemed too inaccessible to many non-methodologists, so we compromised by putting the word "effect" (i.e., total "effect") into quotes everywhere it appears and we remind readers throughout that these data do not support causal inferences. Our goal is to strike a reasonable balance between rigor of terminology and accessibility.

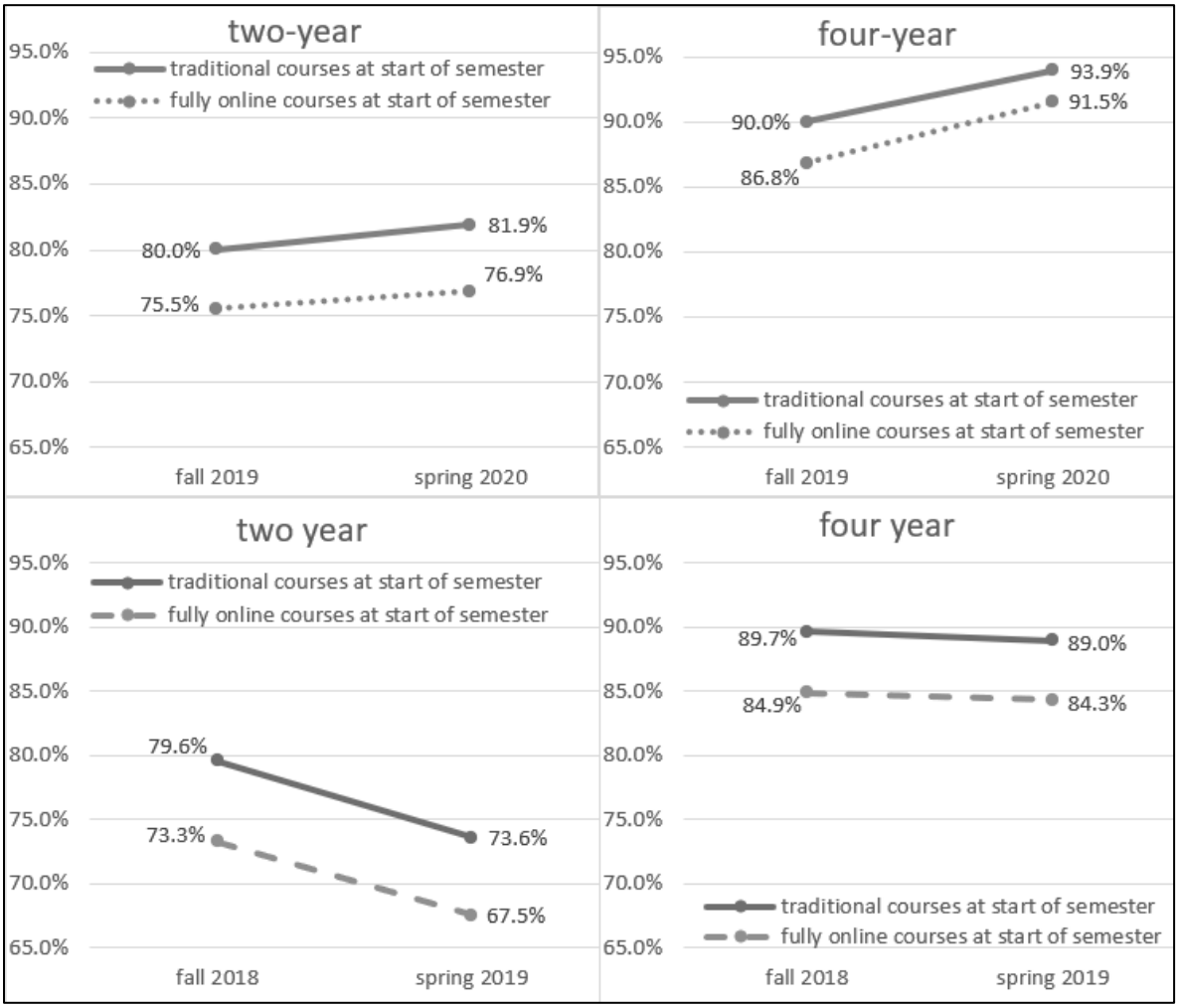
Results and Discussion

General Trends for the Gap in Successful Course Completion by Medium During the Pandemic

Outcomes were compared for fully online versus traditional courses for the same student in fall and spring terms. Figure 2 shows the mean successful course completion rate by student by term, by instructional medium at two- and four-year colleges (patterns from the year 2018–2019 are included for reference in the second row of Figure 2).

Figure 2

Mean Successful Course Completion Rates for Online Versus Traditional Courses Taken by Same Student During Pre-pandemic Fall Versus ERT Pandemic Spring Term



Note. Prior year included in bottom row for comparison.

Students[7] across colleges did significantly better ($p < 0.001$) in both fully online and traditional courses in the pandemic versus pre-pandemic term. The gap between fully online and traditional mode courses remained unchanged—in both terms, at both types of colleges, students were significantly less likely to successfully complete fully online than traditional mode courses ($p < 0.001$), even after traditional courses were moved fully online. The main difference between 2018–2019 and 2019–2020 is that students at both two- and four-year colleges, in both traditional and fully online courses, did better during the pandemic term than expected based on trends from 2018–2019. This may be because instructors were more flexible and lenient during

[7] When we use the term "students," we are referring students in this sample, or those who had enrolled in at least one fully online and one traditional course both terms.

the pandemic term (Altingdag et al., 2021; Chan, 2023; Gelles et al., 2020; Lederman, 2020). However, the *relative gaps* between student performance in fully online versus traditional courses did not change significantly during the pandemic; there was *no significant interaction* between term and course medium in 2019–2020 in multilevel linear probability fixed effects models (see Table 1).

Table 1

Difference in Successful Course Completion (w/C- or Better) within Students, by Term and Course Medium (with Interaction Term)

	Base Model (no covariates)				Full Model (with controls)[a]			
Two-year	Coeff.	SE	*p*	sig.	Coeff.	SE	*p*	sig.
pandemic term	0.026	0.007	0.000	***	0.027	0.007	0.000	***
fully online course	-0.045	0.008	0.000	***	-0.044	0.008	0.000	***
fully online x pandemic term	-0.004	0.012	0.725	*NS*	-0.005	0.012	0.672	*NS*
Four-year	Coeff.	SE	*p*	sig.	Coeff.	SE	*p*	sig.
pandemic term	0.041	0.002	0.000	***	0.037	0.002	0.000	***
fully online course	-0.028	0.003	0.000	***	-0.027	0.003	0.000	***
fully online x pandemic term	0.005	0.004	0.203	*NS*	0.004	0.004	0.271	*NS*

Note. * denotes $p < 0.05$; ** denotes $p < 0.01$; *** denotes $p < 0.001$
Multilevel linear probability fixed effects model coefficients (third level is college, second level is individual student, first level is individual course records); reference groups: pre-pandemic term, traditional mode; separate models were run for two- versus four-year colleges.
[a] Control variables include: gender, race/ethnicity, age, GPA category at the start of the semester, first-semester freshman, median household income of zip code

On average, both traditional and fully online courses had significantly better outcomes in the pandemic versus pre-pandemic term, and fully online courses had significantly worse outcomes than traditional courses both terms, even after traditional courses moved online. Students were less likely to successfully complete courses that they originally chose to take fully online than face-to-face, regardless of whether those courses were subsequently taught fully online or not. The lack of an interaction between course medium and term (i.e., a similar gap during both pre-pandemic and pandemic terms) is consistent with H2 rather than H1. This is because we would have expected the gap to shrink or disappear if H1 were true because regardless of the student's original chosen medium, *all* courses shifted fully online during the pandemic term. Whereas if H2 were true, the factors contributing to differences in course outcome would have been related only to the initially *chosen* medium of the course rather than the actual medium as implemented; thus, in this case we would expect the gap to remain regardless of whether courses were subsequently taught online or not. We now explore each of these hypotheses explicitly and in depth.

Exploring H1 (Course Medium Directly Influences Course Outcomes)

H1 hypothesizes that the online medium itself results in worse outcomes. However, the patterns observed in Figure 2 and Table 1 are not consistent with H1 (i.e., an observed lack of interaction between term and medium). If H1 were true, and the reason for the lower rates of

successful course completion in fully online versus traditional courses pre-pandemic was the online course medium itself, then barring other potential confounding factors[8], we would expect the gap between fully online and traditional courses to shrink, close, or reverse during the pandemic term when traditional courses moved online. Yet, the gap stayed the same; higher education discourse that argues the online medium itself causes worse outcomes is not consistent with patterns observed in this study and bears further examination. This data is observational, and therefore we cannot make causal inferences. We can, however, use other features of the data to explore whether the lack of support for H1 observed in the patterns is or is not reinforced by other features of the data.

Robustness Check: Proportion of Term that was Fully Online as an Instrumental Variable

At CUNY, most colleges switched to ERT five weeks into the 15-week long spring 2020 term; however, three community colleges began the spring term about one week before the switch to ERT. We call the former *less-online* colleges, and the latter, *more-online* colleges. If the online medium itself were responsible for gaps in successful course completion between traditional and fully online courses (H1), then we would expect the gap between traditional and fully online courses to be larger in the less-online colleges (where traditional courses had more in-person instruction) than in the more-online colleges (where traditional courses had almost no in-person instruction). We chose one community college from the less-online colleges and one community college from the more-online colleges; these two colleges were chosen because they had a similar baseline pattern: they had patterns of course outcomes by medium (fully online versus traditional) that were not statistically different ($p = 0.852$) in the fall pre-pandemic semester. We ran a three-way interaction model to determine if there were differences in term-by-online interactions for more-online versus less-online colleges (see Table 2).

[8] To counteract this pattern (the gap between fully online and traditional courses remaining the same during the pandemic term), confounding factors would have to be factors that either: (a) lowered outcomes during the pandemic in fully online courses (those that were originally online) more than in traditional courses (those that were not originally online but moved online during ERT); or (b) raised outcomes during the pandemic in traditional courses (those that were not originally online but moved online during ERT) more than fully online courses (those that were originally online). We discuss potential confounding factors in the final section.

Table 2

Successful Course Completion (with a C- Grade or Better) for Students at Sample "Less-Online" and "More-Online" Colleges[a]

	Base Model (no covariates)				Full Model (with controls)[b]			
	Coeff.	SE	*p*	sig.	Coeff.	SE	*p*	sig.
pandemic term	0.160	0.045	0.000	**	0.148	0.048	0.002	**
fully online course	0.072	0.063	0.256	*NS*	0.074	0.063	0.240	*NS*
fully online x pandemic term	-0.167	0.091	0.065	·	-0.169	0.090	0.059	·
more-online college	0.051	0.058	0.382	*NS*	0.015	0.060	0.803	*NS*
pandemic term x more-online college	-0.077	0.055	0.164	*NS*	-0.071	0.057	0.210	*NS*
fully online x more-online college	-0.025	0.073	0.728	*NS*	-0.018	0.072	0.804	*NS*
fully online x pand. term x more-online	0.117	0.105	0.265	*NS*	0.111	0.104	0.286	*NS*

Notes. · denotes $p < 0.10$; * denotes $p < 0.05$; ** denotes $p < 0.01$; *** denotes $p < 0.001$
Multilevel linear probability model fixed effects model coefficients (second level is individual student, first level is individual course records; reference groups: pre-pandemic term, traditional mode course, less-online college), with three-way interaction by term, course medium, and more-online versus less-online college status.
[a] "More-online" colleges began their term at the end of January (five weeks of the fifteen-week term were taught in-person); "less-online" colleges began their term at the beginning of March (one week of the fifteen-week term was taught in-person).
[b] Control variables include: gender, race/ethnicity, age, GPA category at the start of the semester, first-semester freshman, median household income of zip code.

There was no significant three-way interaction between term, fully online medium, and whether the college was a more- or less-online college, suggesting that the proportion of instruction that was spent online during the pandemic term did not have a significant relationship to course outcomes in this sample.[9] As with prior analysis, results are not consistent with H1 (the online medium itself causes lower rates of successful course completion). If H1 were true, we would expect the proportion of time spent in online instruction during the spring pandemic term to be inversely related to successful course completion: i.e., outcomes in traditional courses at more-online colleges to go down more than at less-online colleges. However, the relationship

[9] We ran the same models with all of the more-online and less-online colleges and did not find significant three-way interactions in any models. We report the results from the model containing one college per group because there was significant variation among the community colleges in terms of baseline outcomes for fully online courses, and we felt that it was most appropriate to report coefficients for those colleges which were most similar at baseline.

between outcomes in traditional versus fully online courses was the same regardless of how many weeks of in-person instruction traditional courses had before the switch to fully online instruction. While these are not causal models and we therefore cannot definitively rule out H1, the patterns observed in this data using timing/duration of online onset as an instrumental variable are not consistent with this hypothesis. Using the difference in number of weeks of in-person instruction before the switch to online courses as a type of instrumental variable[10] is just a different but complementary way of separating out factors that simultaneously impact students' choice of course medium from time spent in the actual course medium (and thus another way of disentangling the mechanisms described by H1 from those described by H2).

Exploring H2 (Type of Course Students Choose to Take Online Predicts Outcomes)

We next considered H2 (differential outcomes online versus face-to-face can be explained in part by students choosing to take different courses online than face-to-face). If H2 were true, in the absence of other confounding factors, we would expect the gap between fully online and traditional courses to remain constant from the pre-pandemic to the pandemic term, since students selected the delivery mode prior to the onset of the pandemic. The chosen delivery mode is in this context a proxy for all the course characteristics that lead a student to choose the online medium for that course (and which may also subsequently drive outcomes in that course as well).

In Figure 2 and Table 1, the gap between fully online and traditional courses was practically identical in both terms, which is consistent with H2 (that gaps between online and face-to-face courses could be explained by students choosing to take different courses in each medium). The pattern observed here is thus consistent with higher education discourse that argues that differential outcomes in online courses are related to student choices to take different courses online versus face-to-face. We contemplate the potential impact of compounding factors in the final section, but first we considered an alternate analysis with different data that provides a robustness check to see if alternate data is also consistent with H2 or not.

Robustness Check: "Importance" of Courses to a Student's Degree Plan as A Predictor of Online Versus Traditional Course Outcomes During ERT

Courses may have various characteristics that make a student more likely to take them online versus face-to-face (or vice versa), and the particular characteristics that are salient may vary by student, making this a complex process. Thus, it is impossible to explicitly categorize all course characteristics that may impact student medium choice; however, we can focus on subsets of these characteristics that have a clear definition. Thus, for this robustness check we focused on one subcategory of characteristics that has been found in the literature to impact students' choice of course medium. There is evidence that students take less "important" courses online and more "important" courses face-to-face (Bejerano, 2008; Jaggars, 2014; Mann & Henneberry, 2014); to test whether this might explain the persistent gap in online versus traditional course outcomes even after traditional courses moved online, we conducted an additional analysis using a subset

[10] Whether a student spent more or less time in online instruction in traditional courses before the pandemic switched all instruction fully online was uncorrelated with either (a) confounding factors (i.e., course characteristics) that determined students' course medium selection or (b) with gaps in course outcomes between online and in-person conditions (outside of the measured potential relationship between time spent in online versus in-person instruction during the pandemic term).

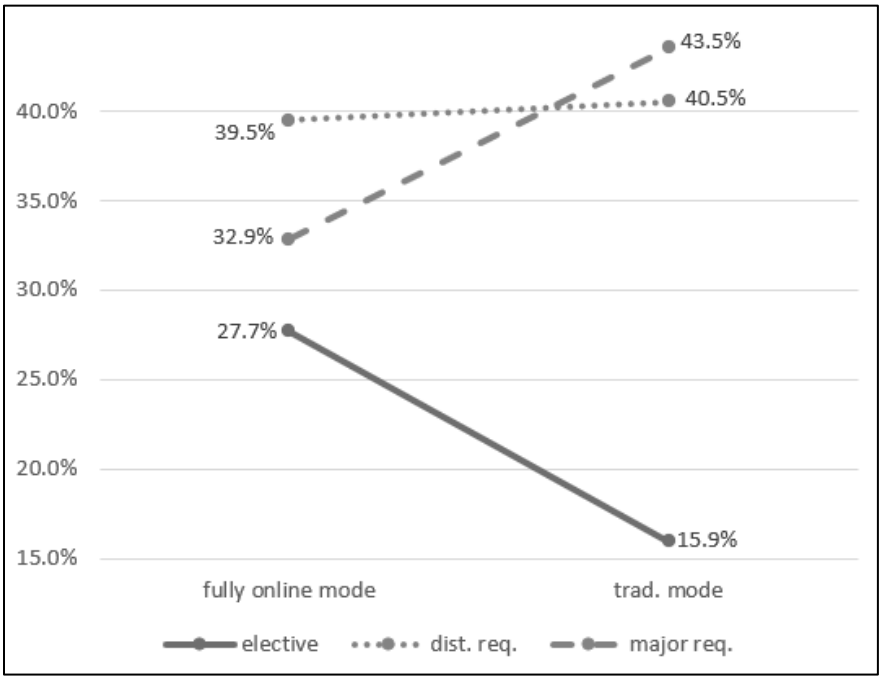
of the data. One measure of the "importance" of a student's course to their degree is whether it counts as a major requirement, distributional requirement, or an elective. These categories have also been shown to explain variation in online versus face-to-face courses in prior research (Wladis et al., 2014, 2017).

For this robustness check, we studied only students enrolled at the largest CUNY community college because this analysis required hand-coding of all courses for each course-major combination in the dataset.[11] We examined the extent to which (a) the courses that students took in the fully online medium were more likely to be elective or distributional requirements rather than major requirements (Figure 3); and (b) the extent to which this distribution of course type by medium may explain the patterns found for course completion by medium by term (Tables 3 and 4).

[11] We excluded students majoring in "liberal arts," which is a broad major not associated with any discipline, intended to help fulfill all distributional requirements for any four-year degree once students transfer to a four-year college.

Figure 3

Proportion of Courses by Type That Students Chose in Fully Online Versus Traditional Mode



Students were significantly more likely to take elective courses fully online ($p < 0.001$) and significantly more likely to take major requirements as traditional courses ($p < 0.001$)[12]; they were roughly equally likely to take distributional requirements as fully online or traditional courses. Courses that were either elective or distributional requirements had significantly lower rates of successful completion than major requirement courses (2.3 percentage points lower, $p = 0.01$; and 4.3 percentage points lower, $p < 0.001$, respectively); thus, we explored the extent to which this may explain the lower rates of completion for courses fully online during the pandemic term by using KHB analysis to run mediation models (Table 3).

Table 3

Mediation Analysis of How Elective Versus Major Requirement Course Type Might Mediate the Relationship Between Course Medium and Successful Course Completion (with a C- or Better)

	Base Model (no covariates)				Full Model (with controls)[a]			
	Coeff.	SE	*p*	sig.	Coeff.	SE	*p*	sig.
Total "effect"	-0.048	0.010	0.000	***	-0.051	0.010	0.000	***
Direct "effect"	-0.035	0.011	0.001	**	-0.040	0.010	0.000	***
Indirect "effect"	-0.014	0.003	0.000	***	-0.011	0.004	0.006	**

Note. * denotes $p < 0.05$; ** denotes $p < 0.01$; *** denotes $p < 0.001$

[12] Significance values come from linear probability models using course type to predict course medium; because of space constraints, we do not report model coefficients here.

In the spring pandemic term for students at the largest two-year college, using KHB method, clustering by student, with and without covariates.
[a] Control variables include gender, race/ethnicity, age, GPA category at the start of the semester, first-semester freshman, median household income of zip code.

Elective-vs-major-requirement-course-type significantly mediated the relationship between course medium and course outcomes during the pandemic term, explaining on average 28% of the difference in course outcomes by medium. In other words, 28% of the outcome gap by course medium can be explained by the fact that students were more likely to take electives than major requirements online and simultaneously had poorer outcomes in elective versus major requirement courses. Thus, the types of courses that students chose to take online explained a significant and substantial part of the difference in outcomes between fully online versus traditional courses during the pandemic (even though it was limited to just one narrow subset of course characteristics), further supporting H2.

This occurred even though the model in Table 3 is only a rough measure of the courses a student may consider "important." This categorization excludes nuanced distinctions that likely go into student decision-making; there are factors that make a course "important" that cannot be captured by our categorization of elective versus major requirement. For example, a student may take courses critical to future career plans but that are technically electives for their current major (e.g., students who plan to transfer to a specific major at a four-year college, or to a graduate program such as medical school, may take classes that are not required by their current major). Because of this, and evidence that students are more likely to take "important" courses in the traditional mode than online (Bejerano, 2008; Jaggars, 2014; Mann & Henneberry, 2014), the models in Table 3 likely *underestimate* the potential power of the "importance" of a course (and other course-level factors) in explaining differences in outcomes for courses that students elect to take online versus face-to-face.

We did not include distributional requirements in mediation models in Table 3. This is because there was an interaction "effect" between distributional requirement categorization and course medium in predicting outcomes (see Table 4). When an interaction term between course type and medium is included in the model, even more of the variation in course success rates are explained, because students successfully complete distributional requirements at significantly lower rates compared to major requirements in the fully online medium than in the traditional mode ($p < 0.001$). Though roughly the same proportion of fully online and traditional courses are distributional requirements, the outcomes in these courses are significantly worse when they are chosen to be taken fully online. One possible explanation is that students might have elected to take less "important" distributional requirements online rather than face-to-face; this is in line with existing research (Bejerano, 2008; Jaggars, 2014; Mann & Henneberry, 2014) and consistent with elective versus major requirement patterns observed in Table 3. The interaction patterns observed in Table 4 are consistent with this explanation. Because of this interaction "effect," we did not include mediation models for distributional versus major requirements in Table 3 (indirect "effect" estimates average over the substantial interaction between course medium and course type in terms of their relationship to outcomes). This provides further evidence that the mediation model in Table 3 likely underestimates the "effect" of course "importance" (and other course-level factors) on differential course outcomes by course medium.

The results in Tables 3 and 4 establish that the importance of a course to a student's degree plan may play a role in explaining differences in outcomes between courses that students elect to take online versus face-to-face. Even though this is not sufficient to make causal conclusions, this finding does reinforce prior research (Jaggars, 2014; Wladis et al., 2014, 2017).

Table 4

Successful Course Completion (with a C- or Better), by Course Medium and Type in the Spring Pandemic Term, for Students at the Largest Two-Year College

	Base Model (no covariates)				Full Model (with controls)[a]			
	Coeff.	SE	*p*	sig.	Coeff.	SE	*p*	sig.
fully online	-0.070	0.006	0.000	***	-0.071	0.006	0.000	***
course type								
elective	-0.023	0.009	0.010	*	-0.022	0.009	0.016	*
dist. req.	-0.043	0.007	0.000	***	-0.037	0.007	0.000	***
fully online	-0.047	0.011	0.000	***	-0.047	0.010	0.000	***
course type								
elective	-0.021	0.012	0.076	·	-0.018	0.012	0.125	·
dist. req.	-0.027	0.009	0.002	**	-0.021	0.009	0.018	*
fully x course type								
fully x elective	-0.016	0.018	0.391		-0.019	0.018	0.303	
fully x dist. req.	-0.051	0.016	0.001	**	-0.052	0.016	0.001	**

Note. · denotes $p < 0.10$; * denotes $p < 0.05$; ** denotes $p < 0.01$; *** denotes $p < 0.001$
Multilevel linear probability fixed effects model coefficients (second level is individual student, first level is individual course records); reference groups: traditional mode, major requirement; two separate models: with (top model) and without (bottom model) interaction term.
[a]Control variables include: gender, race/ethnicity, age, GPA category at the start of the semester, first-semester freshman, median household income of zip code.

Exploring Potential Confounding Relationships

Because this study compared course outcomes within students within terms, student-level variation and term-level variation has been controlled. While different students may have experienced differential impacts of the pandemic, we compared the outcomes in each student's fully online versus traditional courses during the pandemic term to each other, so these impacts would theoretically have influenced both online and traditional courses for that student similarly. Yet, ERT is different from online courses offered outside emergency conditions, so we next considered aspects of potential or theorized instructor/course-level differences during ERT that might have differentially impacted student outcomes in online versus traditional courses during the pandemic term:

C1) ERT courses might have lower quality than fully online courses, because the traditional course instructors had less time to prepare/less experience online (Legon et al., 2020; McDaniel et al., 2020; Stewart, 2021);
C2) Instructors might have been more lenient/flexible in ERT courses than fully online courses. Gelles et al. (2020) and Lederman (2020) found that instructors in general were more lenient during the pandemic without considering original course medium; Altingdag

et al. (2021) found the opposite of the theorized relationship needed to explain existing patterns—instructors teaching courses that were originally fully online were more lenient during the pandemic;

C3) ERT courses might have been more likely to have been offered synchronously (Legon et al., 2020; McDaniel et al., 2020), whereas fully online courses at CUNY (and elsewhere) tend to be asynchronous (Schaffhauser, 2017), and theoretically, outcomes in synchronous courses might be better than in asynchronous ones.

We outline what patterns in the data would be consistent with each hypothesis in isolation. Then, we consider how potential combinations of factors could theoretically lead to different data patterns, considering how these combinations could or could not result in the patterns observed at CUNY during the terms just before/after pandemic onset.

Hypothesis/confounder	**Overall course outcome gap pattern that would be consistent with this factor in isolation**	**Gap change direction**
H1. Lower rates of successful course completion online are caused by the online medium.	Gap *closes* during pandemic term (during ERT, when all courses moved online).	↓
H2. Lower rates of successful course completion online are caused by students selecting "less important" courses for the online medium.	Gap *stays the same* during the pandemic term (because course medium was selected prior to pandemic onset, and the factors that impact medium choice are also what explain poorer outcomes).	→
C1. Traditional courses were of lower quality during pandemic term because of ERT.	Gap *reverses* during pandemic term, as outcomes in traditional courses become worse than those in fully online courses.	↓
C2. Instructors were more flexible/lenient in ERT (traditional courses) than fully online courses during the pandemic term.	Gap *increases* during pandemic term, as students in ERT courses do even better than those in fully online courses due to increased flexibility/leniency.	↑
C3. ERT (traditional) courses were more likely to be synchronous than fully online courses; synchronous outcomes may be better than asynchronous ones.	Gap *stays the same* (or shrinks only slightly) during pandemic term (if synchronicity rather than medium is the reason for previously observed gaps in outcomes between online and face-to-face courses).	→

The gap in successful course completion between fully online and traditional courses stayed the same during the pandemic term, so the data is not consistent with H1, C1, or C2 alone; each would have increased or decreased the gap. For any of these factors to have played a role, they would have had to do so in combination with other factors that exerted an effect on the gap in the opposite direction. We consider each possible combination.

C1 (ERT courses were of lower quality) and H1 (the online course medium itself causes worse outcomes) are the two factors that would each, in isolation, *decrease* the gap between fully

online and traditional course outcomes during the pandemic. Thus, for either of these to have played a role, another factor that *increased* the gap between fully online and traditional course outcomes during the pandemic would have to simultaneously have played a role—of the theorized confounders, only C2 is in the correct direction to produce that pattern. It is possible that the gap remained the same during the pandemic if ERT courses were of lower quality (C1) *and* instructors teaching these courses were *also* more flexible/lenient than instructors of fully online courses (C2); these two effects could "cancel" each other out, producing the patterns observed. Or H1 (the online course medium itself causes worse outcomes) could have played a role if C2 (instructors in ERT courses were more lenient/flexible than instructors of fully online courses)—was enough to override existing negative effects of ERT (C1) in addition to existing negative effects of the online medium itself (H1).

Thus, for C1 and/or H1 to be true, it seems likely that C2 would also have to have exerted an effect. Yet, what evidence exists suggests the opposite of C2—fully online instructors were shown to be more lenient than ERT instructors in spring 2020 (Altingdag et al., 2021; Chan, 2023; Gelles et al., 2020; Lederman, 2020). Thus, this scenario seems unlikely. It is possible that some as-yet unidentified confounding factor exists that increased the gap between courses that were originally fully online and those that were not but moved online during the pandemic, and therefore that C1 and/or H1 could have played some role in course outcomes, but current hypotheses can't account for this.

Further, robustness checks of H1 using the proportion of instruction in traditional courses that occurred fully online during the spring 2020 term as a proxy for exposure to the online medium (the robustness check) were inconsistent with H1: the time that traditional courses spent in the fully online medium during the pandemic term was *not* correlated with significant change in the gap between traditional and fully online courses. While H1 cannot be definitively ruled out, it does not appear to be consistent with the data patterns observed.

The unchanged nature of the gap in successful course completion between traditional and fully online courses during the pandemic term is consistent with H2 (gaps between online and face-to-face courses can be at least partially explained by students taking less "important" courses online) and C3 (ERT courses were more likely to be synchronous; synchronous courses have better outcomes than asynchronous ones). While there is no concrete evidence, the literature does show a tentative connection between synchronous (versus asynchronous) instruction and engagement (e.g., McBrien et al., 2009) and there is research linking engagement to online course outcomes (e.g., Protopsaltis & Baum, 2019), so it is possible that C3 played a role. Since this would have left the gap unchanged, whether C3 played a role should have had no effect on whether or not the evidence supports H1 or H2.

Looking at changes to the gap in course outcomes between fully online and traditional courses from the pre-pandemic fall to pandemic spring terms is alone insufficient to determine if H2 (gaps between online and face-to-face courses can be at least partially explained by the fact that students take less "important" courses online) played a role, even though the data pattern is consistent with this hypothesis. We cannot rule out the possibility that this data pattern was caused by some combination of C2 working simultaneously with H1 and/or C1, even though there is no evidence (in existing research or the data here) that is consistent with that possibility.

None of those data patterns would rule out the potential effect of H2, as all could have occurred in tandem to produce the observed patterns.

Further, careful exploration of the "importance" of courses to students' majors was strongly consistent with H2, suggesting that H2 does play some role in explaining outcomes in fully online versus traditional courses for students who choose to enroll in both mediums. Because students did select more elective courses in the online medium and more major requirements as traditional courses, and because these differences in course type significantly explained differences in course outcomes both during and prior to the pandemic, this provides robust evidence that the kinds of courses that students select for the online medium likely explain, to some extent, differences in outcomes.

Limitations

The CUNY student population is more diverse and urban than the average U.S. college; therefore, while the study results are likely generalizable to a wider national student population, the dataset is not nationally representative. The CUNY data offers an excellent context in which to investigate the impact of the pandemic on traditionally underserved groups; however, it does limit the groups to which results can be generalized.

Given the complexity influencing student lives in spring 2020, we caution against simplistic causal interpretation of the relationships identified here. Rather, this work is useful for considering hypotheses about which factors may lead to student success in online environments in light of observed data. Future work using a research design more appropriate for causal inference is necessary, such as randomized controlled trials that focus on students who choose to take courses online but randomize *which* courses those students take fully online versus face-to-face. Such designs would be difficult to implement but would allow us to assess impacts of the online medium on the students who in practice actually enroll in these courses (while removing confounding effects of course characteristics on both choice of medium and course outcomes).

Researchers should also be prepared to take advantage of system shocks (like COVID-19) that allow data to be collected that are relevant for quasi-experimental analysis designs as have been employed in this study that get closer to investigation of causality than basic regression designs by eliminating various forms of confounding effects. It is important that such studies carefully consider the target population in particular—for example, by considering the impact of these system shocks not just on the whole student population—but on the subset of students who choose to take online courses, since this is the population who will be impacted by policies and structures regarding online course-taking. That this study is observational does not undermine the value of this research, however, it does impose limitations —assertion of causal relationships would be an inappropriate application of these results.

Conclusion and Implications

For students who took both fully online and traditional courses, fully online courses had lower course completion rates both terms, even after traditional courses moved online for ERT after pandemic onset. Differences in outcome by the originally selected course medium did not change even after face-to-face courses shifted fully online. Differences in face-to-face course

outcomes also did not appear to be explained by the proportion of the term spent in in-person versus fully online instruction prior to pandemic onset; data in this study were not consistent with the often-cited hypotheses that the online medium itself leads to worse rates of successful course completion. Because this was not a causal study we cannot definitively rule out the possibility that the online medium itself still plays some role in outcomes, even if no such pattern was visible in these data.

This study also explored whether differences in online versus face-to-face courses could be explained by course characteristics, where the same characteristics that prompt students to take certain courses online also simultaneously result in poorer course outcomes. Courses that were originally fully online and those that were originally not fully online (but shifted fully online during the pandemic) had the same outcomes gap as observed prior to the pandemic, which is consistent with this hypothesis. Further evidence was observed using major requirement versus electives as a proxy of a course's "importance," which has been shown to be a factor in students' choice of medium in the research literature. The data showed that students take electives more often and major requirements less often online; these differences explained a significant portion of the gap between outcomes in courses that were originally fully online versus not-fully-online (before and during the pandemic). This suggests that when differences are observed in outcomes for online versus face-to-face courses, some portion of this difference is likely due to differences in the *types* of courses that students choose to take online versus face-to-face. When students choose to take course types that already have poorer outcomes online instead of face-to-face, the subsequent outcomes in those online courses appear worse on average, even though in these cases online versus in-person instruction has little to do with these differences in outcomes.

While this research is observational, there is currently a lack of causal studies. Thus, our findings may help institutions make policy decisions in the absence of better information. This study suggests that institutions should be cautious about enacting policies based on assumptions that the online medium itself produces worse outcomes for students who elect to take courses online. For instance, it is common practice to employ online readiness surveys (with no comparative for face-to-face courses), which may discourage students from enrolling online even when they are not at risk of dropout (Wladis & Samuels, 2016). Institutions may also enact other restrictions (i.e., who can register, how many online courses) that might reduce access to college for students who need online courses (women, parents, or students who are time-poor [Daymont et al., 2011; Jaggars et al., 2013; Wladis et al., 2018, 2022]) to progress. Additionally, institutions may have put pressure on students to return to taking courses face-to-face (e.g., Bauer-Wolf, 2020; Kumar, 2021); however, for students with disabilities/health issues, long commutes, and/or extensive work/family commitments, taking courses in person instead of online may have negative costs (Daymont et al., 2011; Jaggars et al., 2013; Wladis et al., 2018, 2022). Exerting such pressure can introduce negative stigma for those who need online accommodations. The patterns observed in this study suggest that restricting access to online courses may not produce any measurable benefit, and it has the potential of causing harm by limiting students' access to college, particularly for students who have already struggled historically to gain access to higher education.

Institutions may instead want to focus on supporting the needs of students who elect to take online courses. Policies focused on helping students to successfully complete courses that they consider to be less "important" to their degree may be a promising start. Students who enroll in online courses are more time-poor than their peers who take only courses face-to-face (Wladis et al., 2022), and it may be that when time-strapped students need to drop a class, the least "important" course is the first to be sacrificed (Hachey et al., 2023). Addressing students' time poverty may be one approach for improving outcomes in online courses (and in all their courses more generally). More research is also needed to determine which interventions would best support students in successfully completing elective courses; since students often choose to take these courses online, this could help boost online (as well as face-to-face) outcomes. It is critical to note that this study deliberately investigated students who had *chosen* to take courses online; implications cannot be extended to students who do *not* voluntarily select online courses (e.g., students who are forced online during an emergency). It is possible that forcing students who have *not* selected online courses into a fully online course could have a negative impact on their course outcomes; further research (e.g., well-designed RCT conducted outside of ERT conditions) is necessary to determine what kind of impact this might have. In the absence of such evidence, perhaps the best approach that institutions could take is to respect individual students' choices about which medium is best for them, making sure that there are enough online and face-to-face courses to accommodate all students' course-taking preferences. This includes offering enough sections across each medium to respond to student demand, and also de-stigmatizing student choices by recognizing students' agency and ability to assess what types of learning environments work best for them given their individual resources and constraints. Giving every student equal access to college may require honoring student choices about the particular course medium that works best for them.

Acknowledgements

This research was supported by a grant from the National Science Foundation (Grant No. 1920599). Opinions reflect those of the authors and do not necessarily reflect those of the granting agency.

We would also like to thank Zun Tang and the whole staff at the CUNY Institutional Research offices; Vivek Upadhyay and the staff at the CUNY Registrar's Office; Youngren Ponnuraj and the staff at CUNY Computer Information Services; and Zev Jeremias, CUNY's Executive Director of Technology, for their invaluable help in obtaining the data for this research. We would also like to thank Albert Sun and the *New York Times* for providing help obtaining data on COVID-19 case rates in New York City for this study. In addition, we would like to thank project managers Shristi Karim and Emma Wilson for their work on this project, and Chris Rhoads for his extensive and helpful feedback on the manuscript during its preparation.

References

Alarifi, B. N. & Song, S. (2024). Online vs in-person learning in higher education: Effects on student achievement and recommendations for leadership. *Humanities and Social Science Communications, 11*(86). https://doi.org/10.1057/s41599-023-02590-1

Allison, P. D. (2002). *Missing data*. Sage.

Alpert, W. T., Couch, K. A., & Harmon, O. R. (2016). A randomized assessment of online learning. *American Economic Review, 106*(5), 378–82. https://doi.org/10.1257/aer.p20161057

Altindag, D. T., Filiz, E. S. & Tekin, E. (2021). Is online education working? National Bureau of Economic Research. Working Paper 29113. http://www.nber.org/papers/w29113

Attewell, P., Heil, S., & Reisel, L. (2011). Competing explanations of undergraduate noncompletion. *American Educational Research Journal*, *48*(3), 536–559. https://doi.org/10.3102/0002831210392018

Bauer-Wolf, J. (2020). "Trump administration amps up pressure to reopen colleges." *Education Dive*. https://www.educationdive.com/news/trump-administration-amps-up-pressure-to-reopen-colleges/581166/

Baum, S., & McPherson, M. S. (2019). *The human factor: The promise and limits of online education*. Daedalus. https://www.amacad.org/sites/default/files/publication/downloads/Daedalus_Baum%20McPherson_Fall2019.pdf

Bejerano, A. R. (2008). The genesis and evolution of online degree programs: Who are they for and what have we lost along the way?, *Communication Education*, *57*(3), 408–414. https://doi.org/10.1080/03634520801993697

Bejerano, A. R. (2008). Face-to-face or online instruction? Face-to-face is better. *Communication Currents—National Communications Association.* https://www.natcom.org/communication-currents/face-face-or-online-instruction-face-face-better

Bettinger, E. P., Fox, L., Loeb, S., & Taylor, E. S. (2017). Virtual classrooms: How online college courses affect student success. *American Economic Review, 107*(9), 2855–75. https://doi.org/10.1257/aer.20151193

Bird, K., Castleman, B. L., & Lohner, G. (2020). Negative impacts from the shift to online learning during the COVID-19 crisis: Evidence from a statewide community college system. *EdWorking Papers.* https://doi.org/10.26300/gx68-rq13

Boland, W. C. (2021). *New York: Minority serving institutions in the Empire State*. Rutgers Graduate School of Education— Center for Minority Serving Institutions. https://cmsi.gse.rutgers.edu/sites/defau lt/fles/NewYorkReport.pdf

Bright, K., & Vogler, J. S. (2024). Learning online versus learning in person: A mixed-methods approach to understanding how student preferences and perceptions have evolved since the pandemic. *Online Learning, 28*(4), 56–78. https://doi.org/10.24059/olj.v28i4.4565

Buis, M. L. (2010). Stata tip 87: Interpretation of interactions in nonlinear models. *The Stata Journal*, *10*(2), 305–308. https://journals.sagepub.com/doi/pdf/10.1177/1536867X1001000211

Burns, K., Duncan, M., Sweeney II, D. C., North, J. W. & Ellegood, W. A. (2013). A Longitudinal comparison of course delivery modes of an introductory information systems course and their impact on a subsequent information systems course. *MERLOT Journal of Online Learning and Teaching*, *9*(4), 453–467. https://jolt.merlot.org/vol9no4/burns_1213.pdf

Castro, M. D. B., & Tumibay, G. M. (2021). A literature review: Efficacy of online learning courses for higher education institution using meta-analysis. *Education and Information Technologies, 26,* 1367–1385. https://doi.org/10.1007/s10639-019-10027-z

Chan, H. Y., & Wang, X. (2018). Momentum through course-completion patterns among 2-year college students beginning in STEM: Variations and contributing Factors. *Research in Higher Education, 59*, 704–743. https://doi.org/10.1007/s11162-017-9485-8

Chan, C. K. Y. (2023). A review of the changes in higher education assessment and grading policy during COVID-19. *Assessment & Evaluation in Higher Education, 48*(6), 874–887. https://doi-org.utep.idm.oclc.org/10.1080/02602938.2022.2140780

Daymont, T., Blau, G., & Campbell, D. (2011). Deciding between traditional and online formats: Exploring the role of learning advantages, flexibility, and compensatory adaptation. *Journal of Behavioral and Applied Management*, *12*(2), 156. https://www.academia.edu/28848297/Deciding_Between_Traditional_and_Online_Formats_Exploring_the_Role_of_Learning_Advantages_Flexibility_and_Compensatory_Adaptation

Doo, M. Y., Zhu, M., & Bonk, C. J. (2023). A systematic review of the research topics in online learning during COVID-19: Documenting the sudden shift. *Online Learning, 27*(1), 15–45. https://doi.org/10.24059/olj.v27i1.3405

Faidley, J. (2018). *Comparison of learning outcomes from online and face-to-face accounting courses*. (Publication No. 3434) [Doctoral dissertation, East Tennessee State University] *Electronic Theses and Dissertations*.. https://dc.etsu.edu/etd/3434

Figlio, D., Rush, M., & Yin, L. (2013). Is it live or is it Internet? Experimental estimates of the effects of online instruction on student learning. *Journal of Labor Economics*, *31*(4), 763–784. https://doi.org/10.1086/669930

Fischer, C., Xu, D., Rodriguez, F., Denaro, K. & Warschauer, M. (2020). Effects of course modality in summer session: Enrollment patterns and student performance in face-to-face

and online classes. *The Internet and Higher Education, 45,* 100710. https://doi.org/10.1016/j.iheduc.2019.100710

Fischer, C., Baker, R., Li, Q., Orona, G.A., & Warschauer, M. (2021). Increasing success in higher education: The relationships of online course taking with college completion and time-to-degree. *Educational Evaluation and Policy Analysis, 44*(3). https://doi.org/10.3102/01623737211055768

Francis, M. K., Wormington, S. V., & Hulleman, C. (2019). The costs of online learning: Examining differences in motivation and academic outcomes in online and face-to-face community college developmental mathematics courses. *Frontiers in Psychology, 10*, Article 2054, 1–12. https://dx.doi.org/10.3389%2Ffpsyg.2019.02054

Garman, D. E. (2012). Student success in face-to-face and online sections of biology courses at a community college in east Tennessee. (3515024). *Electronic Theses and Dissertations.* Paper 1408. https://dc.etsu.edu/etd/1408

Gelles, L. A., Lord, S. M., Hoople, G., Chen, D. A., & Mejia, J. A. (2020). Compassionate flexibility and self-discipline: Student adaptation to emergency remote teaching in an integrated engineering energy course during COVID-19. *Education Sciences, 10*(11), 304. https://doi.org/10.3390/educsci10110304

Green J., Hanckel, B., Petticrew M., Paparini, S., & Shaw, S. (2022). Case study research and causal inference. *BMC Medical Research Methodology, 22*(1), 307. https://doi.org/10.1186/s12874-022-01790-8

Hachey, A. C., Conway, K. M., Wladis, C., & Karim, S. (2022). Post-secondary online learning in the U.S.: An integrative review of the literature on undergraduate student characteristics. *Journal of Computing in Higher Education, 34*, 708–768.

https://doi.org/10.1007/s12528-022-09319-0

Hachey, A. C., Wladis, C., & Conway, K. M. (2023). Investigating online versus face-to-face course dropout: Why do students say are they leaving? *Educational Science, 13*(11), 1122. https://doi.org/10.3390/educsci13111122

Hart, C. M., Friedmann, E. A., & Hill, M. (2018). Online course-taking and student outcomes in California community colleges. *Education Finance and Policy, 13*(1), 42–71. https://doi.org/10.1162/edfp_a_00218

Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching and online learning. *Educause Review*, *27*, 1–12. https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning

Hyllegard, D., Deng, H., & Hunter, C. (2008). Why do students leave online courses? Attrition in community college distance learning courses. *International Journal of Instructional Media, (35),*4, 429–434.

https://link.gale.com/apps/doc/A273359032/AONE?u=anon~1f708b5a&sid=googleScholar&xid=33d1b979

Jaggars, S. S. (2014). Choosing between online and face-to-face courses: Community college student voices. *American Journal of Distance Education*, *28*(1), 27–38. https://doi.org/10.1080/08923647.2014.867697

Jaggars, S. S. & Bailey, T. (2010). Effectiveness of fully online courses for college students: Response to a department of education meta-analysis. *Community College Research Center (CCRC), Teachers College Columbia University*. https://files.eric.ed.gov/fulltext/ED512274.pdf

Jaggars, S. S. (2014). Choosing between online and face-to-face courses: Community College student voices. *American Journal of Distance Education, 28*(1), 27–38. https://doi.org/10.1080/08923647.2014.867697

Jaggars, S. S., Edgecombe, N., & Stacey, G. W. (2013). What we know about online course outcomes. Research overview. *Community College Research Center (CCRC), Teachers College Columbia University*. https://ccrc.tc.columbia.edu/media/k2/attachments/what-we-know-about-online-course-outcomes.pdf

Jaggars, S. S., & Xu, D. (2010). Online learning in the Virginia Community College System. *Community College Research Center (CCRC), Teachers College Columbia University*. https://ccrc.tc.columbia.edu/media/k2/attachments/online-learning-virginia.pdf

Jaggars, S. S., & Xu, D. (2014). How do online course design features influence student performance? *Computers & Education, 95,* 270–284. https://doi.org/10.1016/j.compedu.2016.01.014

James, S., Swan, K., & Daston, C. (2016). Retention, progression and the taking of online courses. *Online Learning*, *20*(2), 75–96. https://files.eric.ed.gov/fulltext/EJ1105922.pdf

Johnson, H. P., & Mejia, M. C. (2014). *Online learning and student outcomes in California's community colleges*. Public Policy Institute. https://www.ppic.org/content/pubs/report/R_514HJR.pdf

Young, S., & Bruce, M. A. (2011). Classroom community and student engagement in online courses. *MERLOT Journal of Online Learning and Teaching, 7*(2), 219–230. https://jolt.merlot.org/vol7no2/young_0611.pdf

Kebritchi, M., Lipschuetz, A., & Santiague, L. (2017). Issues and challenges for teaching successful online courses in higher education: A literature review. *Journal of Educational Technology Systems*, *46*(1), 4–29. https://doi.org/10.1177/0047239516661713

Keramidas, C. G. (2012). Are undergraduates ready for online learning? A comparison of online and face-to-face sections of a course. *Rural Special Education Quarterly, 31*(4), 25–32. https://doi.org/10.1177%2F875687051203100405

Kofeoed, M. S., Gebhart, L., Dilmore, D., & Moschitto, R. (2021). Zooming to class? Experimental evidence on college students' online learning during COVID-19. *IZA Institute of Labor Economics.* https://ftp.iza.org/dp14356.pdf

Kohler, U., Karlson, K. B., & Holm, A. (2011). Comparing coefficients of nested nonlinear probability models. *The Stata Journal*, *11*(3), 420–438. https://doi.org/10.1177/1536867X1101100306

Kumar, D. (2021). At the University of Florida, a rise in face-to-face classes prompts pushback. *Tampa Bay Times*. https://www.tampabay.com/news/education/2021/02/06/at-university-of-florida-a-rise-in-face-to-face-classes-prompts-pushback/

Lederman, D. (2020). How teaching changed in the (forced) shift to remote learning. *Inside Higher Education*, *22*. https://www.insidehighered.com/digital-learning/article/2020/04/22/how-professors-changed-their-teaching-springs-shift-remote

Legon, R., Garrett, R., Fredericksen, E. E., & Simunich, B. (2020). CHLOE 5: The pivot to remote teaching in spring 2020 and its impact. *Quality Matters & Eduventures Survey of Chief Online Officers,* 1–37. https://qualitymatters.org/qa-resources/resource-center/articles-resources/CHLOE-project

Manly, C. A., & Wells, R. S. (2015). Reporting the use of multiple imputation for missing data in higher education research. *Research in Higher Education*, *56*(4), 397–409. http://doi.org/10.1007/s11162-014-9344-9

Mann, J. T., & Henneberry, S. R. (2014). Online versus face-to-face: Students' preferences for college course attributes. *Journal of Agricultural and Applied Economics, 46*(1), 1–19.

McBrien, J. L., Cheng, R., & Jones, P. (2009). Virtual spaces: Employing a synchronous online classroom to facilitate student engagement in online learning. *International Review of Research in Open and Distributed Learning*, *10*(3). https://doi.org/10.19173/irrodl.v10i3.605

McDaniel, C., Suffern, C., Joo, J., & Alamuddin, R. (2020). *Student and faculty experiences with emergency remote learning in spring 2020. ITHASK S+R Research Report.* https://doi.org/10.18665/sr.314276

McDonough, C., Roberts, R. P., & Hummel, J. (2014). Online learning: Outcomes and satisfaction among underprepared students in an upper-level psychology course. *Online Journal of Distance Learning Administration*, *17*(3). https://www.westga.edu/~distance/ojdla/fall173/mcdonough_roberts_hummel173.html

McDougall, A., Orlov G., & McKee, D. (2020). Learning during the COVID-19 pandemic. https://voxeu.org/article/learning-during-covid-19-pandemic

Nguyen, T. (2015). The effectiveness of online learning: Beyond no significant difference and future horizons. *MERLOT Journal of Online Learning and Teaching, 11*(2), 309–313. https://jolt.merlot.org/Vol11no2/Nguyen_0615.pdf

Nora, A., & Snyder, B. P. (2009). Technology and higher education: The impact of e-learning approaches on student academic achievement, perceptions and persistence. *Journal of College Student Retention: Research, Theory & Practice, 10*(1), 3–19. https://doi.org/10.2190/CS.10.1.b

Norton, E. C., & Dowd, B. E. (2018). Log odds and the interpretation of logit models. *Health Services Research*, *53*(2), 859–878. https://doi.org/10.1111/1475-6773.12712

Pao, T. C. (2016). Nontraditional student risk factors and gender as predictors for enrollment in college distance education. *Chapman University Digital Commons—Educational Studies Dissertations.* https://doi.org/10.36837/chapman.000009

Patterson, B., & McFadden, C. (2009). Attrition in online and campus degree programs. *Online Journal of Distance Learning Administration, 12*(2). http://www.westga.edu/~distance/ojdla/summer122/patterson112.html

Protopsaltis, S., & Baum, S. (2019). *Does online education live up to its promise? A look at the evidence and implications for federal policy*. George Mason University Center for Education Policy and Evaluation. https://jesperbalslev.dk/wp-content/uploads/2020/09/OnlineEd.pdf

Schaffhauser, D. (2017). Asynch delivery and the LMS still dominate for online programs. *Campus Technology.* https://campustechnology.com/articles/2017/05/22/asynch-delivery-and-the-lms-still-dominate-for-online-programs.aspx

Shea, P., & Bidjerano, T. (2014). Does online learning impede degree completion? A national study of community college students. *Computers & Education*, *75*, 103–111. https://doi.org/10.1016/j.compedu.2014.02.009

Smith, B. (2010). E-learning technologies: A comparative study of adult learners enrolled on blended and online campuses engaging in a virtual classroom. *Capella University ProQuest Dissertations & Theses, 4313143* https://search.proquest.com/docview/746605945

Smith, N. D. (2016). *Examining the effects of college courses on student outcomes using a joint nearest neighbor matching procedure on a state-wide university system.* [Doctoral dissertation, North Carolina State University]. North Carolina State Repository. https://repository.lib.ncsu.edu/server/api/core/bitstreams/e35a2133-8235-4ff7-b1f5-4a591ba8f32e/content

Stewart, W. H. (2021). A global crash-course in teaching and learning online: A thematic review of the empirical Emergency Remote Teaching (ERT) studies in higher education during Year 1 of COVID 19. *Open Praxis, 13*(1), 89–102. https://search.informit.org/doi/pdf/10.3316/informit.758902304536019

Tanyel, F., & Griffin, J. (2014). A ten-year comparison of outcomes and persistence rates in online versus face-to-face courses. *B>Quest* (2014). https://www.westga.edu/~bquest/2014/onlinecourses2014.pdf

Tsang, J. T. Y., So, M. K. P., Chong, A. C. Y., Lam, B. S. Y., & Chu, A. M. Y. (2021). Higher education during the pandemic: The predictive factors of learning effectiveness in COVID-19 online learning. *Education Sciences, 11*(8), 446. https://doi.org/10.3390/educsci11080446

Wladis, C. W., Conway, K. M., & Hachey, A. C. (2014). The role of enrollment choice in online education: Course selection rationale and course difficulty as factors affecting retention. *Journal of Asynchronous Learning Networks, 18(3).* *https://files.eric.ed.gov/fulltext/EJ1043163.pdf*

Wladis, C. W., Hachey, A. C., & Conway, K. M. (2015). Which STEM majors enroll in online courses and why should we care? The impact of ethnicity, gender, and non-traditional student characteristics. *Computers & Education, 87,* 285–308. https://doi.org/10.1016/j.compedu.2015.06.010

Wladis, C., & Samuels, J. (2016). Do online readiness surveys do what they claim? Validity, reliability, and subsequent student enrollment decisions. *Computers & Education*, *98*, 39–56. https://doi.org/10.1016/j.compedu.2016.03.001

Wladis, C. W., Conway, K. M., & Hachey, A. C. (2017). Using course-level factors as predictors of online course outcomes: A multilevel analysis at an urban community college. *Studies in Higher Education. 42*(1), 1–17. https://doi.org/10.1080/03075079.2015.1045478

Wladis, C., Hachey, A. C. & Conway, K. M. (2018). No time for college? An investigation of time poverty and parenthood. *Journal of Higher Education*. https://doi.org/10.1080/00221546.2018.1442983

Wladis, C., Hachey, A. C., Conway, K. M. (2022). Time poverty: A hidden factor connecting online enrollment and college outcomes? *Journal of Higher Education, 94*(5), 609–637. https://doi.org/10.1080/00221546.2022.2138385

Wladis, C. W., Hachey, A. C., & Conway, K. M. (2023). Did emergency remote teaching and the COVID-19 pandemic exacerbate inequities? Considering institution type, gender and race/ethnicity. *The Journal of Higher Education, 95*(3), 313–349. https://doi.org/10.1080/00221546.2023.2250696

Wladis, C. W., Hachey, A. C. & Conway, K. M. (2024). It's about time: The inequitable distribution of time as a resource for college, by gender and race/ethnicity. *Research in Higher Education, 65*(7)*,* 1614–1646. https://doi.org/10.1007/s11162-024-09796-5

Xu, D., & Jaggars, S. S. (2014). Performance gaps between online and face-to-face courses: Differences across types of students and academic subject areas, *The Journal of Higher Education, 85*(5), 633–659. https://doi.org/10.1080/00221546.2014.11777343

Xu, D., & Jaggars, S. S. (2013). Adaptability to online learning: Differences across types of students and academic subject areas. *Community College Research Center (CCRC), Teachers College Columbia University*. https://ccrc.tc.columbia.edu/media/k2/attachments/adaptability-to-online-learning.pdf