ABSTRACT

LEVINSON, NATHANAEL SIMEON. The Effect of Delivery Method on Student Success in General Chemistry Courses in Community College (Under the direction of Dr. Michelle Bartlett).

This study was conducted to elucidate the effect of delivery methods on student success in general chemistry courses. A propensity score model was used to evaluate the effect of delivery method on student success within a general chemistry course, as measured by letter grade distribution, GPA average, and grade differential from a first semester course to a second semester course. The demographics between students in online chemistry courses and in-person chemistry courses were virtually identical, as were the grade distributions for both semesters of general chemistry. Age and Pell status were found to be statistically different between the two groups, but the effect size was very small. The average grade for both delivery methods was a B for the first semester course, and for both delivery methods the average change in grade from first to second semester courses was a drop of approximately one letter grade. Overall, the difference in outcomes between delivery methods was negligible, indicating that, on average, online chemistry courses provide an academically similar environment for students.

The Effect of Delivery Method on Student Success in General Chemistry Courses in Community College

by
Nathanael Levinson

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Education

Community College Leadership

Raleigh, North Carolina

2024

APPROVED BY:

Dr. Michelle Bartlett Committee Chair	Dr. James Bartlett, II
Dr. Carrol Warren	Dr. Donna Petherbridge

DEDICATION

To my LORD and my God, King of the universe, to His only Son, my savior, who took my sin and gave me His righteousness, and to the Holy Spirit, who enlightens my eyes and delights my soul, and shows me the paths of righteousness.

To my family: Papa, who reminded me that fortune flavors the bowl; Mama, who proved you're never too old to start playing DnD; Hannah, from where she encourages me to keep learning; Ethan, the tru greetest and my oldest and closest friend; Micah, who is indeed the fastest and the nicest coffee snob you'll ever meet; Joel, the biggest shrimp I know and a fellow Yugioh enjoyer; Abigail, who got all the artistic talent yet remains Kingus Bingus Dingus; Naomi, fellow breadmaker, who can now say I'm giving Doctor Levinson.

To my brothers in Christ who prayed for me always: Greg, who laughs with me as we watch terrible movies; Daniel, who is literally a 10-year older version of me; Tim, in spite of all the times his bad tile-drawing luck got us killed; Matt, who probably still owes me a cigar and a bottle of whiskey.

This would not have been possible without all your advice, encouragement, friendship, and prayer. I love you all.

BIOGRAPHY

Nathanael Levinson was born and raised in Charlotte, North Carolina. He got bachelor's degrees in both chemistry and biology at Vanderbilt University in Nashville, then his master's in chemistry at Georgia Tech in Atlanta. He is now back in Charlotte, teaching chemistry at Central Piedmont Community College. He is an avid chef and gamer, has his own cooking YouTube channel, The Kitchen Chemist, and a podcast where he waxes idiotic about film and television, Dearly Debated.

ACKNOWLEDGMENTS

Thank you to my chair, Dr. Michelle Bartlett, who guided me through the dissertation process and was readily available in spite of a leaky roof. Thanks also to her husband and my committee member Dr. James Bartlett, who got angry on my behalf and helped demystify propensity score matching (although he bullied me into using it). Thanks also to my other committee members, Dr. Carrol Warren and Dr. Donna Petherbridge, who were patient with me and made themselves available despite their busy schedules.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 - Introduction

Introduction

Online learning has become more accessible than ever before, thanks to technological improvements (Ives, 2006), widespread adoption (Cox, 2005), and increasing ease of economic access (Oakley, 2010). For many students, online offerings provide flexible learning opportunities that allow them to pursue their educational goals without warping their lives around a scholastic schedule (Moloney, 2010). There is, however, a debate on whether online courses offer similar instructional levels as traditional, in-person courses (Carr, 2000). Data indicates that in general, drop rates are higher for online courses than for in-person courses (Bettinger, 2017). There is also a question of whether a student who has taken an introductory course online is as well prepared for advanced courses as their in-person counterparts. The COVID-19 pandemic threw these questions into the foreground, as many institutions moved temporarily to an online-only format (Bird, 2022; Castleman, 2022; Kolack, 2020).

Chemistry courses are a core educational component for students who are either pursuing a degree in physical sciences or in a medical field, such as nursing (Coley, 1973). Because there is an incredibly high volume of students engaged in introductory chemistry courses (Coley, 1973), it is worthwhile to determine whether online chemistry courses show the same efficacy as in-person courses with regards to student success. The available data for comparing online and in-person courses is generic and groups courses that have vastly different subjects and goals, so there is merit in examining the data in a more granular fashion to see if there is a notable difference in the efficacy of a chemistry course based on its delivery method.

This dissertation will explore the student success outcomes of various course delivery methods for introductory chemistry courses. The delivery modes examined will be traditional in-person, online, hybrid, and blended type delivery. The study will be a quantitative analysis of student success factors, including course completion data, degree completion data, and transcript-level grade data. The data will be obtained from Large Generic Community College and the NCCCS dashboard. The goal will be to establish whether the course delivery method had a statistically significant effect on student performance, both within the course and in subsequent chemistry courses.

Background of the Problem

The development of the internet, along with the rapid pace of technological advancement, have thrust us into a digital age (Bowen, 2015; Brown, 1996). Even in its earliest forms, the internet gave educators a glimpse into the evolution of correspondence courses: online distance education. The preponderance of distance education had grown significantly over the years, with most universities and community colleges offering online options for students (Brown, 2001; Kumar, 2017).

Increased access to online resources, along with faster internet speeds and the widespread use of smartphones have made distance education available to virtually anyone (Saykali, 2019). These online courses cover every conceivable subject, up to and including laboratory sciences. General chemistry is an introductory course that must be taken by a great number of students outside of the major field: biology, nursing, and dental programs to name a few all require at least one semester of chemistry, and in North Carolina alone, around ten thousand students enroll in the first semester of general chemistry every year (NCCCS, 2023).

Statement of the Problem

Research has shown that students' attrition rates are higher in online courses, and completion of the course is less likely than for in-person courses (Gregori, 2018; Jordan, 2015).

In fact, drop-out rates for online students can be as high as 50% (Prendergast, 2003). Online courses are a different experience from traditional in-person courses (Redmond, 2018; Vonderwell, 2003), requiring more autonomy in learning than in-person courses.

Community colleges, being open institutions, face even greater challenges when it comes to online learning. Community colleges serve broader swaths of the population and are a key component of education for students who can be considered non-traditional (Goldrick-Rab, 2010). Additionally, community colleges generally have a substantial vocational aspect, which also requires unique approaches to online learning (Bragg, 2001). Students who drop out of online courses report feeling isolated and overwhelmed (Bambara, 2009). A student who drops out of a course experiences a set-back in degree completion, with significantly reduced chances of finishing the degree (McKinney, 2019). The attainment of an Associate's degree, as well as technical certificates, is tied to higher earnings and generally more positive outcomes, especially for people from minority backgrounds (Belfield, 2011). Therefore, it is crucial to assess the effects of delivery methods on chemistry courses as it may have a significant impact on students' futures.

Generally, the data indicates that online courses have significantly higher dropout rates than in-person courses, but the data is not generally discipline specific (Carr, 2000; Hobson, 2018; Levy, 2007). The 'hard' sciences are very different from social sciences or liberal arts because they generally focus on objective physical realities, whereas social sciences and liberal arts are more subjective, being open to interpretation and opinion (Van Gigch, 2002), best viewed as a distinction between mechanistic–analytical domains on the one hand, and behavioral–biological domains, on the other. In educational terms, speaking from personal experience as an educator of more than 10 years, this means that in physical science courses, generally there will be one or more correct answers to a problem, while all answers outside this relatively small domain are considered incorrect. In social sciences and liberal arts, while there may be what can be considered a 'best' answer, there is room for subjectivity and opinion, and there are scenarios where there is no wrong answer or failure state. This is not to say that there is no overlap in successful instructional methods (Foerster, 2003), simply that the outcomes are more or less dependent on different factors (Gabrielle, 2003). Given these differences, there is merit in obtaining discipline-specific data for online vs in-person courses.

There has not been a specific study examining student success in general chemistry courses based on modality, according to my search of the available literature. As online classes are on the rise, it is imperative to examine whether an online chemistry course is academically similar to a traditional, in-person chemistry course.

Purpose of the Study

The purpose of this study is to determine whether there exists a statistically significant difference in completion, as measured by course and program completion rates, in taking general chemistry courses online versus in-person. This study will elucidate what impact the delivery method of a general chemistry course has on students' success. This will be measured for success within the course and success in more advanced chemistry courses. The data garnered from this study will help chemistry faculty begin to assess how they approach building and delivering courses based on delivery methods and will fill a gap in the research pertaining to online education.

Theoretical Framework

As this work is focused specifically on online learning and in-person learning for adults in a community college setting, a framework of andragogy will be utilized. Additionally, Bean and Metzner's (1985) model of student attrition will be utilized to explain the quantitative results. This framework is germane to this study as it focuses on the unique challenges of adult learners and non-traditional students, which comprise the bulk of the community college population (Philibert, 2008).

Andragogy

Andragogy refers broadly to an instructional methodology targeted at adult learners. It is a marginal point of debate whether andragogy constitutes a branch of pedagogy or is a separate educational construct (Akin, 2014). Andragogy is a broad term, originally developed as a theory of learning by M.S. Knowles in the 1970s, and generally refers to a more active approach to learning, where the student is more involved in self-learning and bears a similar level of responsibility to the teacher (Zmeyov, 1998). Due to the nature of community colleges, the majority of students fall under the "nontraditional" category (Philibert, 2008). Andragogy is a better framework for both nontraditional students and for online classes, both major considerations in this research, and is therefore an appropriate framework for this study.

Andragogy vs Pedagogy

In general, traditional pedagogical methods focus more heavily on rote and memorization, with the onus of the transmission of information on the teacher, while andragogy focuses primarily on six principles which separate it: that adults need to understand why they are learning, that adults learn primarily through trial-and-error, that adults take personal responsibility for what they choose to learn, that adults prefer practical matters over theoretical concerns, that adults prefer problem-based environs to content-based environs, and that adults lean more heavily on intrinsic motivators (Knowles, 1977).

Andragogy in Online Education

The inherent nature of distance education situates itself more in line with andragogy than traditional pedagogy, and even though distance education does not necessarily involve adult learners, the fundamental principles of andragogy still apply (Chametzky, 2014). Therefore, the theories of andragogy suit this research as a framework for analyzing online courses.

Bean and Metzner's Model of Student Attrition

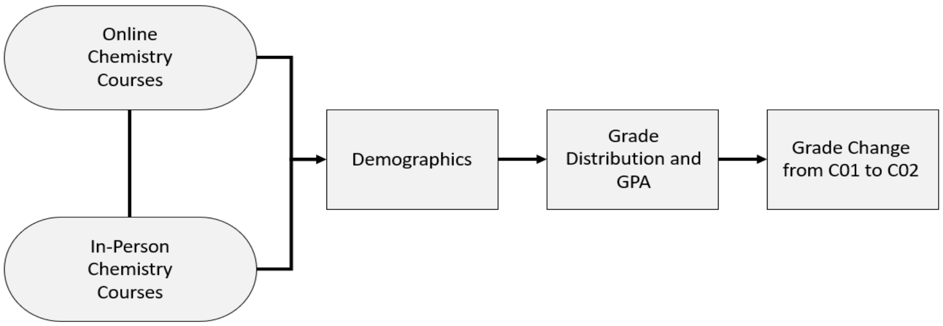
Additionally, there are several well-attested models of student persistence which, in conjunction with the above-mentioned models, will aid in assessing the qualitative aspects of completion rates in both in-person and online chemistry courses. The primary framework here is Bean and Metzner's (1985) model of student attrition, which assesses various external and internal driving factors for nontraditional students such as age, educational goals, satisfaction, and family responsibilities in determining reasons for students dropping out of a program.

Conceptual Framework

The conceptual framework shown in Figure 1.1 visualizes the factors from both online and in-person chemistry courses that will be assessed in this study. Demographic data will be used to assess differences in populations that may exist between delivery methods. Class-level data in the form of grade distribution will be used to compare individual course success across delivery methods. Additionally, the average grade change from C01 to C02 will be examined to determine if delivery method in earlier courses impacts outcomes in later courses.

Figure 1.1

Conceptual Framework of the Study



In order to assess differences between modes of course delivery, demographic differences will be assessed, followed by grade distribution and GPA assessments, and finally how the delivery method of C01 affects C02 outcomes.

Research Questions

1. What are the demographics that constitute in-person vs online chemistry courses of students who enrolled in consecutive semesters of general chemistry in community college?
2. Is there a statistically significant difference in grade distribution based on delivery method among students who enrolled in consecutive semesters of general chemistry in community college?
    a. What are the grade distributions for online general chemistry courses?
    b. What are the grade distributions for in-person general chemistry courses chemistry courses?
    c. Is there a statistically significant difference between a and b?

3. Does the delivery method of the first semester of general chemistry impact the outcome of the second semester of general chemistry? This question involves the following:

    a. What are the GPA (as defined by A, B, C, D, and F) differences between first- and second- semester general chemistry for online chemistry courses?

    b. What are the GPA (as defined by A, B, C, D, and F) differences between first- and second- semester general chemistry for in-person chemistry courses?

    c. Is there a statistically significant difference between a and b?

    d. Does changing the delivery method from first to second semester affect grade outcomes?

Significance

Chemistry is a fundamental subject that many students outside the major field will end up taking; as demonstrated by Oliver et al (2021), where 25% of the total student population was enrolled in introductory chemistry courses, with 20% of those non-STEM majors. Due to the importance of the general coursework, along with the increasing prevalence of online offerings, it is important to establish whether online chemistry is as effective in educational terms as in-person classes. If it is not, then the gaps must be identified and addressed. If it is, then any perceived stigma of online courses should be summarily dealt with. In either case, it will definitively establish the role of online coursework in general chemistry. This is important as most of the data and research regarding the relative effectiveness of online courses is not discipline-specific, but rather considers all of an institution’s courses in general. Due to the inherently different nature of the content and assessment in varying courses, for example, English and math, taking such a broad approach means that nuance slips through the gaps. To this end, this study will fill the particular gap in knowledge around success in chemistry courses based on delivery methods, as well as the effectiveness of online delivery methods in preparing students for more advanced courses. It will serve as a first step in elucidating the differences between online and in-person courses for chemistry students and will reveal any gaps in student success based on delivery method in a course-specific manner. This information can then be used as a springboard for future studies on what aspects of courses might be particular stumbling blocks and what perceptions exist regarding course delivery methods. This study may serve as a cornerstone for the optimization of online chemistry courses.

Definition of Terms

Course performance: Relative achievement in an individual class, as measured by letter grade (Picciano, 2012)

Degree completion: The successful attainment of an Associate's or Bachelor's degree (Oseguera, 2009)

Hybrid/Blended: Exact definition varies based on institution-specific guidelines, but involves distinct portions of a class being exclusively online while maintaining a more infrequent class meeting (Helms, 2010)

In-person/Face-to-Face (f2f): A course in which the students and faculty are present in the same room for lectures and exams, a traditional course. (Helms, 2010)

Online: A course in which the students never meet with the faculty in a physical location but exclusively access lectures and exams online, may be synchronous or asynchronous. (Helms, 2010)

Limitations

The limitations of the study are the number of students data is available for, the detail of data available, and the years data is available for. As data is being pulled from a single university, although that university has a large student population there are still only a few hundred students who take chemistry each semester. Additionally, the university only provided data for students who took the first and second semesters of general chemistry in consecutive semesters, with data pulled from Fall 2021 to Spring 2023.

Delimitations

This study will focus on the chemistry program at a single university, and the data will not be disaggregated. Utilizing a single university as a data source will minimize the potential variances due to content differences such as textbooks, labs, and online systems. There may be some variance due to the effect of individual teachers, but in order to obtain as broad a sample size as possible, the individual teacher will not be taken into account. While there may be merit in disaggregating the data to assess possible educational gaps, this is outside of the intended scope of the project and will therefore not be done.

Chapter Summary

This chapter laid out the groundwork for the study, emphasizing the importance of filling the knowledge gaps in regard to the differences in efficacy between online and in-person chemistry courses. The theoretical framework of andragogy was outlined and a conceptual framework for the study was provided. The specific research questions that the study will address were succinctly presented.

Organization of the Study

Chapter two covers the relevant literature from the two key areas: chemistry education and distance education. The focus is primarily on the development of distance education and the challenges and opportunities that online education offers. There is also a discussion of andragogy and other adult learning theories. Chapter three details the methodology of the study and the application of statistics to answer the research questions. Chapter four discusses the results of the study, contextualized by the theories utilized. Chapter five presents conclusions, implications for theory and practice, and recommendations for future studies.

Chapter 2 - Literature Review

General chemistry has long been perceived as a very difficult or gatekeeper course, one which may determine whether the student goes on to finish the degree or chooses a different educational path (Carter, 1989). While other STEM courses, such as physics, biology, and calculus, could also be viewed as gatekeepers, the unique blend of novel concepts and applied algebra make chemistry a nightmare for many students (Sirhan, 2007; Sjöström, 2014). Additionally, general chemistry is a laboratory course, often one of the first laboratory courses students will encounter. So, on top of tackling a difficult lecture course, students must also learn how to do laboratory work (Reid, 2007). General chemistry is also a math-heavy subject, with problem-solving beyond mere algebra being heavily incorporated. Many students struggle with the mathematical application of novel concepts, as expressed in word problems.

Since this difficult course is so foundational, it is important to ensure that factors leading to success are codified so that the course offering can be optimized for student success. Nowadays, online offerings are commonplace, even for courses with laboratory components, and these offerings come with their own set of unique advantages and drawbacks. A primary advantage of online courses is the possibility of asynchronicity, allowing students to fit the class into any schedule. Additionally, online components of courses have been shown to potentially aid with student understanding of more difficult concepts (Akçay, 2006; Barak, 2004; He, 2012).

While research has been done in relation to gaps between online and face-to-face (f2f) classes, even in STEM areas, there is little research looking specifically at general chemistry courses and whether online offerings are as effective in teaching these foundational concepts as f2f offerings. This research seeks to assess whether there exists a gap between online and f2f chemistry courses in terms of student success – both in the course and in the degree as a whole – and student performance[NL2] .

History of Online Learning

Online learning is a relatively novel mode of education, as the internet itself is relatively novel in terms of use and distribution. The earliest instances of online distance education from the late 80's and early 90's were referred to as "computer-mediated communication", as a natural segue from mail-based correspondence courses (Harasim, 1990). These early online offerings were fraught with the user-unfriendly designs of the early computer era, requiring knowledge of command lines at a basic level, but the promise of asynchronicity and ready access to materials, as well as the potential for distanced interactivity, led educators to pursue better systems for online education (Mason, 1998). Online education first began to see mainstream appeal in higher education in the early 2000's, and by 2005 over 50% of higher education institutions offered online courses to some extent (Allen, 2003; 2004; 2005). From 2010 to 2020, there was a steady increase in online enrollment, even though from 2012 onwards overall enrollment in higher education began to decline. As of 2018, over six million students are taking at least one distance course (Seaman, 2018). The decreasing costs and flexibility of online courses render them somewhat of an inevitability, and it is crucial to ensure that the online experience is at least equivalent to the in-person experience as more and more students make the shift to online learning.

The Rise of MOOCs

In the 2010's, a variant of online offering saw more mainstream appeal, the Massive Open Online Course (MOOC). MOOCs had been around since the 2000's, but in 2012, only 2.6% of schools offered MOOCs (Allen 2013). MOOCs were seen to have several drawbacks, including the poor perception of the credentials obtained through MOOCs. Despite this, some universities created initiatives centered around MOOCs, such as MIT and Harvard's joint venture edX. An assessment of edX's first offering indicated no major differences between students who took the MOOC and students who took the course in-person (Breslow 2013). As the research questions in this dissertation are not specific to MOOCs (but do not necessarily exclude them), the author recommends Littlejohn and Hood's *Reconceptualising Learning in the Digital Age* (Littlejohn 2018) for a more comprehensive examination of MOOCs. The courses examined in this study would not be considered MOOCs.

Challenges of online learning

The online learning environment presents unique challenges and opportunities for educators when compared to in-person offerings (Larreamendy-Joerns, 2006). Online learning intrinsically allows for asynchronous education, and offers a significant increase in flexibility, as students can perform their work at irregular times as their schedules demand. This brings its own issues, as asynchronous study has issues that synchronous online and f2f courses avoid. One major downside to online learning is the lack of instructor contact, especially in asynchronous courses, as well as the lack of contact with other students, which can lead to a host of emotional and psychological issues for students (Kennette, 2015). The nature of online learning requires heightened independence and responsibility from students, which can be a double-edged sword, especially for students who are unaccustomed to self-teaching. Additionally, online courses suffer far more from technical issues than in-person courses, due to the multiplicity of individual variables in the form of student hardware and internet connections, institution- and instructor-side IT issues, and more (Broadbent, 2015). There also exists a unique opportunity for unethical student behavior (McGee, 2013). Despite the challenges, however, many students prefer the flexibility and availability that distance education offers (Jaggars, 2014). Altogether, online learning, while unlikely to completely supplant in-person education, is now a vital aspect of the course offerings of many institutions, meriting in-depth research into its effectiveness, particularly in light of the COVID-19 pandemic in 2020-2021 that forced many institutions to move completely online for a time, strategies around shifting to a more online and hybrid environment have been increasingly relevant.

Online versus in-person general education

There has been a good deal of research exploring whether online courses offer students similar learning experiences and success rates as traditional in-person offerings (Means, 2012).

The data seems to indicate that students in online courses tend to fare worse than their in-person counterparts (Wavle, 2019). In general, dropout rates from online courses are higher than those of in-person courses and tend to disproportionately disfavor lower achieving students (Bettinger 2017). In fact, there is an intrinsic equity gap that exists in online learning, due to the technological necessities of both hardware and adequate internet connectivity, but this topic is outside the scope of this research.

There have been attempts to determine relevant factors which are involved in student success online, but the issue has proven to be complex, with few clear answers (Garrison, 2000; Harrell, 2008; Kaufmann, 2015; Yukselturk, 2007). There are certain self-study skills that may serve as possible indicators of student success, but only weak associations were found between those strategies and student success (DeTure, 2004; Broadbent, 2015; Kauffman, 2015).

As previously mentioned, online offerings are continually increasing in number and are sought after by students for their flexibility. The resolution of this dilemma is vital to institutional success in the online sphere. The difficulty lies in determining if the gaps in success are due to institution-side factors, such as content delivery and course access, or student-side, such as time management or focus. The delineation of these factors is outside the scope of this study but seems to be a crucial point of understanding for the optimization of online offerings.

Historical Overview of Chemistry Education

Chemistry has historically been considered by students to be a difficult subject (Woldeamanuel, 2014). Since the 1960's, efforts have been made to enrich the delivery of the content in a way that makes it more accessible to students. In the late 90's and early 00's, there was a substantial deal of research into novel methodologies for chemistry instruction, including a focus on the psychology of students , a shift towards application and away from traditional logic-based teaching, and alternative assessment methods to facilitate cognitive skills (Johnstone, 2000; Lloyd, 2010; Mahaffy, 2004; Mahaffy, 2018; Reid, 2000; Talanquer, 2007; Wu, 2004; Zoller 2001).

The issue has been approached from many angles, such as an investigation into the presuppositions that students bring with regards to chemical concepts and vocabulary, to emphasizing the importance of practical demonstrations, and pushing for applied pedagogical research (Bucat, 2004; Bulte, 2006; Cooper, 2018; King, 2012; Meyer, 2003; Mulford, 2002). Today, all other considerations aside, there are two major camps of thought with regards to general chemistry education: the atom-first approach and the matter-first approach, both with their advantages and disadvantages (Talanquer 2013). Recent educational research has led to widespread implementation of the "flipped classroom", including within chemistry education, although there is some debate as to whether this methodology is ultimately beneficial (Seery, 2015; Stöhr, 2020). Anecdotally, the author, who has taught chemistry at a community college for five years, has had some success applying the Socratic method of questioning to help students grasp concepts, but has found that, in alignment with some of the research, that it is the mathematical application of concepts that stymies most students. The author considers the atom-first approach to have a more gradual introduction to these concepts and is therefore theoretically more optimal for first-year students. Although the research shows that practical demonstrations are useful in aiding student knowledge, these can be exceedingly difficult to implement in many classrooms, and the flipped classroom approach is difficult to balance with the need to introduce novel chemistry concepts. Ultimately, chemistry is a difficult subject and there is no clear consensus on the most effective method of teaching it.

The Importance of Online and Technological Components in Traditional Chemistry Education

One consistent facet of chemistry education is the increasing utilization of online and technological components to aid and facilitate teaching chemical concepts. Online homework systems are ubiquitous and have been shown to increase student understanding of subject matter. Research has been done to develop interactive simulations, tools to assess student metaknowledge, and even efforts to gamify chemistry. These efforts are separate from the design and distribution of online chemistry courses; most of these efforts are intended to enhance in-person learning for students (Antunes, 2012; Chittleborough, 2007; Eilks, 2010; Gasrcia-Martinez, 2010; Grove, 2007; Moore, 2014; Parker, 2012 ; Rastegarpour, 2012; Singhal, 2012; Williams, 2011). Online laboratory simulations, while seemingly convenient, do not afford students the opportunity to fail in unique ways, even the best simulation cannot account for every mistake that is possible in an experiment. Additionally, the author can report that his students who took online labs felt disengaged from the purpose of labs, that is, to give a visible demonstration of the concepts learned in lecture, and lab for them became an exercise in data copying. Laboratory simulations may have more utility as preparations for in-person lab, getting the student acquainted with the procedure before they formally conduct the experiment in person (Dalgarno, 2009).

Online versus In-Person Chemistry Education

There is a relative dearth of chemistry-specific research on online versus f2f courses, however there is a decent amount of research regarding general STEM courses, including laboratory courses (Bettinger et al, 2017; Bettinger & Loeb, 2017; Ogot, 2003; Sarvary, 2022). The laboratory component is one aspect that makes chemistry and similar STEM courses a vastly different experience when moved online, as laboratory work is an intrinsically tactile experience, one that is essentially negated when put in the online sphere, even given the increase in quality simulation software. While general research (as previously mentioned) indicated an overall decrease in student success in online courses, this may not be the case in science courses (Murphy 2015), and online students may even show better predispositions towards science education (Murphy, 2015; Perera, 2017). While there are many resources available with regards to the substance of online chemistry courses (Eichler, 2013; Leontyev 2013, Tawfik 2017, Tatli 2010), there is little specific research examining student success in online chemistry courses as compared to traditional in-person courses, a gap this research intends to fill.

It is difficult to cohere the data on general online courses with Murphy and Perera's findings, but it seems likely that students enrolled in a science course volitionally tend to be more technologically focused. It should be noted that the difficulty of maintaining student integrity in online courses is well-documented (McGee, 2013), so it is possible that grade estimates for online courses are artificially inflated. This would mean that the true grade gap between online and in-person is even larger than expected, but this matter is beyond the scope of this research. A reasonable explanation for the apparent synchronicity of online and in-person courses in the sciences may lie in the fact that the hard sciences tend towards quantitative assessments, where other courses may rely primarily on qualitative assessments, i.e. quizzes versus essays. In any case, the purpose of this study is to help elucidate the possible gap between online and traditional courses specifically for chemistry.

Studies have shown that in-person chemistry classes provide opportunities for hands-on learning, real-time feedback, and social interaction, which can enhance student engagement and motivation (Bunce, Flens, & Neiles, 2010; Kuo, Walker, Schroder, & Belland, 2014). In-person classes also allow instructors to customize their teaching to students' needs and respond to their questions and concerns in real-time (Bunce et al., 2010).

However, online chemistry classes can also be effective, especially when designed with interactive and engaging materials that facilitate student learning and promote active participation (Alkhateeb, 2018; Khalil & Ebadi, 2014). Online classes can provide students with more flexibility and control over their learning, as they can access course materials at their own pace and convenience (Khalil & Ebadi, 2014). Additionally, online classes can be more cost-effective and accessible for students who live far from campus or have other constraints that prevent them from attending in-person classes (Alkhateeb, 2018).

Bean and Metzner's Model of Student Attrition in the Literature

Bean and Metzner's Model of Student Attrition was developed to specifically address the dropout rates of nontraditional students (Bean & Metzner, 1985). It does this by assessing 'background' variables such as age, enrollment status, and residence; educational goals, high school performance, and a host of other variables in order to determine what factors impact a nontraditional student's decision to leave an institution. They also consider academic variables such as study skills, academic advising, and course availability, environmental variables such as housing and employment, and social integration variables, with the goal of creating a holistic snapshot of a nontraditional student and the factors that most impact their decision to drop out. This model was developed as a framework for future research, and it has been utilized frequently to that end. For example, Bergman (2014) used the model as a foundation for studying factors impacting adult student persistence, finding that educational goals seemed to have the greatest impact on adult learner retention. Radovan (2019), examined the current state of retention models and found Bean and Metzner's research to still be relevant, especially in light of the ubiquity of online courses. On the other hand, a recent longitudinal analysis by Scheunemann (2022). Which focused specifically on dysfunctional student behavior (i.e. procrastination) as a factor impacting persistence, which is included in Bean and Metnzer's model, found that the data did not align well for any historical theoretical framework for the specific area of student dysfunction. Despite this, the model is still widely utilized and considered as a solid foundation for assessing nontraditional student attrition.

Andragogy as a Theoretical Framework

Andragogy is an adult learning theory developed in the 1970s by MS Knowles (1978), and focuses on the shift in responsibility away from the teacher and onto the student, creating a more active learner. There are two central tenets to Knowles' theory, namely that adult learners are self-directed and autonomous, and second that the teacher acts more as a "facilitator of learning" rather than a "presenter of content" (Reischmann, 2004).

Andragogy is a learning theory that suggests that adults have unique learning needs and preferences that should be taken into consideration when designing educational programs (Knowles, Holton III, & Swanson, 2015). Malcolm Knowles, the pioneer of andragogy, argued that adults are self-directed learners who are motivated by relevance and problem-solving, and that adult learning should be experiential and draw on learners' prior knowledge and experience (Knowles et al., 2015). Merriam and Bierema (2014) provided an overview of the major theories of adult learning, including andragogy. They explored the different ways in which adults learn and the implications for teaching and designing educational programs. They also emphasized the role of motivation and self-directed learning in adult education. Brookfield (2017) offered a critical analysis of andragogy and other adult learning theories. He argued that while andragogy has been influential, it has also been oversimplified and can be overly prescriptive. Brookfield advocated for a more nuanced approach to adult learning that takes into account the diversity of adult learners and their varied learning needs and preferences. Overall, the literature on andragogy emphasizes the importance of understanding the unique characteristics of adult learners and tailoring educational programs to meet their needs (Knowles et al., 2015; Merriam & Bierema, 2014). It also highlights the importance of motivation and self-directed learning in adult education (Knowles et al., 2015; Merriam & Bierema, 2014).

Andragogy has been fairly well represented in the literature, and although the theory has its critics (Hartree, 2006), it is still used as a theoretical framework in education research and has been used in several hundred separate research papers (Henschke, 2011). Anrdagogy has specifically been adapted to distance and online education, as this style of education demands more self-sufficiency of its students. Andragogy has been successfully used as the basis for designing and developing online educational courses and has been called a “fundamental principle of online education” and purportedly achieves the “pinnacle of Bloom’s Taxonomy pyramid” in the online learning environment (Decelle, 2016; Chametzky, 2014; Rossman, 2013; Sato, 2017; Mews, 2020). Despite the large amount of research available on andragogy, there are two major gaps which this research will help fill. First, there is little research comparing online and in-person courses through an andragogical framework. Second, there is little research available specifically for chemistry courses within an andragogical framework. The primary concerns of andragogy, that is the self-direction and autonomy, seems an apt fit for community college students which, as mentioned previously, are largely non-traditional learners. Especially in online courses, students are required to learn on their own to some degree in order to succeed. Unlike a traditional primary education class, community college students are in class for relatively short periods of time relative to the amount of work expected of them. This of course is further reduced in blended/hybrid courses, and for asynchronous online courses, the student is entirely for learning the material. Andragogy, whatever its possible lackings, seems a substantially appropriate fit.

Other Adult Learning Theories

There are numerous other adult learning theories, which have been summarized well by other authors. Arghode et al (2017) provides concise explanations for four other adult learning theories: behaviorism, where learning is tied to responses to external stimuli; cognitivism, which focuses on how the mind processes and retrieves information; constructivism, which holds that knowledge is a psychosocial construct; and humanism, which “assumes the ultimate purpose of learning is to facilitate a self-actualized, autonomous person”. Mukhalalati and Taylor (2019) similarly covers a variety of adult learning theories, including several of the same ones, but arranges them into four different categories: instrumental learning theories, humanistic or facilitative learning theories, transformative learning theories, and social theories of learning.

Andragogy Theory

This theory was developed by Malcolm Knowles and is based on the idea that adults learn differently from children. Andragogy emphasizes self-directed learning and the importance of practical experience. It follows a set of five assumptions of adult learners, as laid out by Knowles:

1. Their self-concept moves from one of being a dependent personality towards being a self-directed human being
2. They accumulate a growing reservoir of experience that becomes an increasingly rich resource for learning
3. Their readiness to learn becomes oriented increasingly towards the developmental tasks of their social roles
4. Their time perspective changes from one of postponed application of knowledge to imme-diacy of application, and, accordingly, their orientation towards learning shifts from one of subject-centeredness to one of performance-centeredness.
5. Their time perspective changes from one of postponed application of knowledge to immediacy of application, and, accordingly, their orientation towards learning shifts from one of subject-centeredness to one of performance-centeredness. (Knowles, 1984)

This theory is most frequently applied to college-age and older students, as opposed to pedagogy, which focuses on K-12. Because of this, it is suited and frequently used for nontraditional students, of whom a subset is students older than the typical college age (Mews, 2020; Wozniak, 2020). One in-depth study using andragogy as a focal point was Ladwig's (2021) examination of online learning in graduate health programs. Ladwig aimed to assess the transition from in-person to online learning that occurred during the COVID-19 pandemic and used andragogy as a framework. Ladwig focused specifically on synchronous versus asynchronous online delivery and found that there were significant performance differences between the two delivery methods.

Transformational Learning Theory

This theory, developed by Jack Mezirow, focuses on how adults learn through transformative experiences. It suggests that adults undergo a process of critical reflection and perspective transformation to make sense of their experiences and develop new ways of thinking. Alam (2022) utilized this theory to examine the development of sustainable education, noting that the adaptable nature of the theory made it ideal for optimizing practice in broad fields. Critteden (2019) writes on the utility of transformational learning theory with regards to digital learning and technology, which is being constantly updated and demands rapid user adaptation.

Experiential Learning Theory

This theory, developed by David Kolb, emphasizes the importance of learning through direct experience. It suggests that adults learn best when they are actively engaged in the learning process and can reflect on their experiences. Kolb developed the theory in the 1970's, building off of work by John Dewey, initially put forward in a series of papers, but later published as a book in 1984, updated in 2014 (Kolb, 2014). Kolb describes learning as a cycle between personal development, work, and education. Kolb also evaluates learning styles and developed a graph to assess an individual's learning strengths. The axes consist of experience vs abstract conceptualization on the y-axis and doing vs watching on the x-axis. In this way, four quadrants are created to give 4 distinct learning styles. The Doing/Feeling quadrant is dynamic learners, or accomodators, who understand concepts sensually rather than logically and seek opportunities to apply problem solving skills to novel dilemmas. The Watching/Feeling quadrant is imaginative learners, or divergers, who use imagination to connect old and new information to assess the importance of study and tend to focus on the big picture. The Watching/Thinking quadrant is analytic learners, or assimilators, who focus on facts and benefit most from practical knowledge. The Doing/Thinking quadrant is common sense learners, or convergers, who focus on application and practical knowledge use. Ferreira (2020) utilized the theory to explore transitions from hybrid to full-time entrepreneurship and maintains that the theory helps provide more nuanced insights into the relevant factors. Moseley (2020), used the theory as a foundation for examining the development of environmental educational programs, focusing specifically on the theory's utility in field components, as those are inherently a practical experience. Towns (2001) specifically focused on applying experiential learning to chemistry as "a sound theoretical basis for expanding activities beyond traditional lecture" (p.1116).

Self-Directed Learning Theory

This theory, developed by Malcolm Knowles, emphasizes the importance of adults taking responsibility for their own learning. It suggests that adults are motivated to learn when they have control over the learning process and can set their own goals. This theory was built on by Alan Tough (1971), who focused on informal learning for adults, stating that "almost everyone undertakes at least one or two major learning efforts a year". Tough performed a series of interviews to assess why adults undertake these projects and found that the most compelling reason for most respondents was the desire to perform an action or master a skill at a higher level in pursuit of a specific goal. Sawatasky (2017) states that self-directed learning has three major components:

1. The process of learning

2. The personality characteristics of the learner
3. Factors of the learner's context that influence self-directed learning

As distance learning inherently involves a degree of self-study, many studies focus on that area, such as Geng (2019). The motivations behind self-directed learning and the efficacy of self-study have become sharply relevant in light of the COVID-19 pandemic, which forced many students into an online environment (Coros, 2021; Wahyudi, 2021). Wahyudi examined how teachers' perceptions of proctoring self-directed learning mismatched with actual student experience, indicating that, despite the name, this theory demands a more hands-on approach in its implementation.

Social Learning Theory

This theory, developed by Albert Bandura, emphasizes the importance of learning through observation and imitation of others. It suggests that adults learn best when they have role models and opportunities to practice new skills in a supportive environment. Bandura focused on disabusing the notion that actions arise purely from internal motivators such as personal desires, and explains aggression as a response to external factors, specifically social and cultural contexts (Bandura, 1977). Bandura posits that children learn behavior by observing adults as role models, and that behavior can be reinforced either through direct communication from the model (i.e., praise or chastisement). He further suggests that behaviors can also be influenced by the observation of consequences applied to the model (i.e., model behaves badly, is subsequently punished, child internalizes avoidance of negative behavior). This process is referred to as encoding. Children were found to be more likely to copy behaviors from models they perceived as being similar to themselves. This theory is most frequently used in assessing negative behaviors, such as unethical or immoral actions (Powers, 2020; Smith, 2021), but this theory has also been applied to education, as when Ahn (2020) used social learning theory as a foundation to examine how role-models affect student outcomes. Ahn synthesized social learning theory with a number of learning processes, such as attentional and motivational processes, to develop a more holistic framework for role-model learning, noting:

> The extant role model literature in the context of education covers a wide range of topics in depth, such as modeling mechanisms and outcomes, but tends to neglect crucial social-cognitive processes that prove useful in providing further insights. (p.8)

Adult Development Theory

This theory, also known as the constructive-developmental theory of adult learning, was developed by Robert Kegan and focuses on the stages of adult development and how these stages affect learning. It suggests that adults go through different stages of cognitive development and that their learning needs change as they progress through these stages. Kegan (1982) posits six developmental stages:

1. In the first stage, the self cannot be separated from the objects being interacted with by the self, and ends once this distinction can be made.
2. In the second stage, the self can not be differentiated from its impulses and has no concept of a "me" and lacks goals or agendas.
3. In the third stage, the self can differentiate impulses from themselves, they "have" desires, rather than "being" those desires. At this point, the individual is "imperial" and does not form relationships based on trust or reciprocity.
4. In the fourth stage, the self becomes social, beliefs and sense of self are now influenced and largely determined by external relationships.

5. In the fifth stage, termed, the self-authored mind, the individual derives its sense of self from internal determinants, with opinions and beliefs arising from within rather than from without, usually accompanied with greater emotional intelligence and control.
6. In the sixth and final stage, self is created through the exploration of the individual's own identity, and a willingness to entertain alternative perspectives

The first two stages are oftentimes combined, as they apply only to very young children and babies. Most uses of this framework for adult learning begin with either the social mind (stage 3).

This theory is most frequently applied to adults in professional careers and professional development, rather than in educational contexts, such as Velardi (2020) examining knowledge exchange between maple syrup producers and beekeepers. Likewise, Lewin (2019) uses the theory to examine personal and professional development in medicine careers.

Situated Learning Theory

This theory, developed by Jean Lave and Etienne Wenger, emphasizes the importance of learning in a social context. It suggests that adults learn best when they are actively engaged in a community of practice and can learn from others with similar interests and experiences. The theory hinges on a process called 'legitimate peripheral participation', which maintains that "learners inevitably participate in communities of practitioners and that the mastery of knowledge and skill requires newcomers to move toward full participation in the sociocultural practices of a community" (Lave & Wenger, 1991, p.29). Choi (2021) summarizes the theory by stating "...learning cannot be achieved if separated from the context in which it occurs. The main assumptions of [social learning theory] are that knowledge must be learned in an authentic context of how it might be used and is created when authentic activity is involved." (p.2)

This theory is most commonly used in educational programs which contain a substantial practical element, such as business (Yeoman, 2019), and construction (Wang, 2020). It has also been applied more broadly, for instance Longmore (2021) used the theory as a foundation for creating a listening protocol for a variety of situations. Choi (2021) utilized social learning theory to help implement conflict-resolution measures in a training program for nursing students and saw a promising reduction in conflict using an intervention with social learning theory as the framework.

Propensity Score Matching in Literature

Propensity score matching is a quasi-experimental quantitative statistical methodology which aims to minimize the effects of selection bias by mimicking full randomization in the sample (Austin, 2011). By scoring data sets on certain variables, data points are compared such that the treated and untreated groups are matched with other data points with similar "propensity scores" (Guo, 2015). This technique is most commonly applied in situations where random sampling is not possible, where there are a large number of confounding variables, or when sample sizes are small (Lane, 2012). Benedetto (2018) lists a number of advantages to the propensity-score matching method. First, all confounding variables are consolidated into a single score, allowing the investigator to include all potential confounding variables. Benedetto notes that this is especially useful when a sample has a small size but a large number of confounders. Second, PSM matches to allow for an 'apples-to-apples' comparison, which helps identify areas of non-overlap between subjects. Third, PSM eliminates the assumption of linearity between the propensity score and the outcomes. Finally, PSM mirrors a randomized method, mitigating the effects of selection bias.

The primary weakness of PSM is that it inherently reduces sample size, as if a subject cannot be matched, it is discarded. Because of this, although PSM is very useful for smaller sample sizes, unaccounted for confounders may result in the discarding of a substantial subset of the data, reducing the impact of the findings (Benedetto, 2018). Heinrich (2012) summarizes thus:

> The matching estimator will not necessarily work in all circumstances; specific conditions have to be met to produce valid impact estimates. First, if the condition requiring one to find untreated units that are similar in all relevant characteristics to treated units is to be satisfied, it is clear that these characteristics must be observable to the researcher. In other words, PSM requires selection on observables; the inability of the researcher to measure one or more relevant characteristics that determine the selection process results in biased estimations of the impact of the program. Second, in order to assign a comparison unit to each treated unit, the probability of finding an untreated unit for each value of X must be positive. (p. 14)

This methodology has been used frequently in educational contexts, to control for factors such as age and socioeconomic status. For example, Li (2019) utilized propensity-score matching to assess whether parental help with homework improves student outcomes. More relevantly to this study, Paulsen (2020) used propensity-score matching to assess learning disparities in online learning, finding that the gap between online and in-person learning was less than previously reported once other variables were accounted for through this methodology. A useful model of how propensity score matching can be used in educational research is provided by Lane (2012). Lane uses a hypothetical scenario involving reading proficiency to demonstrate that propensity score matching is appropriate when demographic covariates may obscure or skew the results.

Conclusion of Literature Review

Online learning is a growing facet of education that is recognized as a valid substitute for traditional in-person teaching. While there is plentiful data generally comparing online and f2f offerings, there is a gap of reliable data when it comes to chemistry courses. The intent of the author is to identify if such a gap exists, why it exists, and possible solutions to mediate the gap. It is important to note that at time of writing, the country is embroiled in the global COVID pandemic. As such, there has been a harsh skew towards online delivery as traditional f2f interaction has been limited by quarantine procedures. The vast majority of education is being done virtually, although a gradual shift back towards in-person learning has begun. This phenomenon will generate aberrant data with regards to online and in-person courses and will likely be excluded from consideration. Additionally, it remains to be seen how the impact of COVID will alter the educational landscape, as it forced many students online who would otherwise never have chosen to do distance education. The author is unwilling to speculate and will therefore be largely ignoring COVID and its impact in the course of their research.

Chapter Summary

This chapter summarizes the relevant literature with respect to chemistry education, distance education, and andragogy as an adult learning theory. Distance education provides unique challenges and opportunities for students, and research seems to indicate that in general distance education options have higher dropout rates. There is a gap in the literature regarding specifics about chemistry courses and also the efficacy of online courses in leading to degree completion for chemistry students which this study hopes to address.

Chapter 3 - Methodology

Chapter Introduction

This chapter details the research design, data and methods used to address the research questions. First, the research design is described, followed by descriptions of the measurements, participants, and instrumentation. The chapter concludes with descriptions of the data collection and data analysis, as well as ethical considerations.

Methodology

The purpose of this study is to examine the differences between course and degree completion, between in-person and online chemistry courses. Further, this study aims to examine if a relationship exists between course success and degree completion. To address this inquiry, the following research questions (RQs) were posed as a guiding framework for the study:

1. What are the demographics that constitute in-person vs online chemistry courses of students who enrolled in consecutive semesters of general chemistry in community college?
2. Is there a statistically significant difference in grade distribution based on delivery method among students who enrolled in consecutive semesters of general chemistry in community college?
    a. What are the grade distributions for online general chemistry courses?
    b. What are the grade distributions for in-person general chemistry courses chemistry courses?
    c. Is there a statistically significant difference between a and b?
3. Does the delivery method of the first semester of general chemistry impact the outcome of the second semester of general chemistry? This question involves the following:

a. What are the GPA (as defined by A, B, C, D, and F) differences between first- and second- semester general chemistry for online chemistry courses?

b. What are the GPA (as defined by A, B, C, D, and F) differences between first- and second- semester general chemistry for in-person chemistry courses?

c. Is there a statistically significant difference between a and b?

d. Does changing the delivery method from first to second semester affect grade outcomes?

Theoretical Framework

The theoretical framework of andragogy was chosen for this study, as it focuses on adult learners and non-traditional learners, which comprise the bulk of community college enrollment (Philibert, 2008). Andragogy is particularly suited to a study examining online education, as the necessities of self-led learning emphasize themselves in an online setting (Chametsky, 2014). Andragogy outlines six principles which guide adult learners: first, that adults seek to understand why they are learning what they are learning; second, that trial-and-error is a major component of adult learning; third, that practical concerns tend to trump theoretical matters; fourth, that adult learners have personal responsibility for the courses they choose; fifth, that problem-based teaching is preferred to content-based teaching; sixth, that intrinsic motivators are more important than extrinsic motivators (Knowles, 1977). The inherent self-paced and self-guided components in online education make andragogical considerations more impactful in that context as compared to a traditional in-person context (Chametsky, 2014), therefore it may be intuited that the demographic construction of the online student population may differ from that of the in-person student population and that the reasons for choosing online or in-person may also depend on these demographics. For this reason, propensity score-matching was used to ensure that any difference in demographics that may exist would be accounted for and the selection bias would be reduced.

Research Design

This study utilizes a quantitative propensity score matching research design. The study is centered around statistical analyses and propensity-score matching to analyze data obtained from a large community college in North Carolina, henceforth referred to as Large Generic Community College. This data includes demographics and transcript data, as well as delivery methods for students who took the general chemistry courses at Large Generic Community College, C01 and C02, in consecutive semesters. The data was analyzed to assess variances between delivery methods (in-person and online) and the results were not disaggregated by demographics for the purposes of this research.

A quantitative analysis is appropriate for this topic, as we are interested in the discrete differences between delivery methods, as measured by the previously detailed metrics, all of which are quantifiable, to assess whether statistically significant differences exist between delivery methods, as measured by the previously mentioned quantifiable metrics. Since our sample is not randomized, and the demographic covariates may be confounding, propensity score matching is an appropriate methodology to apply to the study (Lane, 2012)

Propensity Score Matching Overview

Propensity score matching is a statistical technique used to balance the distribution of covariates between treatment and control groups in observational studies (Austin, 2011). The goal of propensity score matching is to create a pseudo-randomized experiment by matching treated and untreated individuals who have similar values on a set of covariates that are associated with the outcome of interest (Rubin, 2001). The process of propensity score matching involves estimating the propensity score, which is the probability of receiving the treatment based on observed covariates. This score is then used to match treated and untreated individuals using various matching algorithms, such as nearest-neighbor matching or optimal matching (Austin, 2011).

Propensity score matching has been widely used in a variety of fields, including healthcare, education, and social sciences, to address the problem of selection bias in observational studies (Guo & Fraser, 2015; Rosenbaum, 2002). In education research, propensity score matching has been used to evaluate the effectiveness of various interventions, such as tutoring programs (Tackett & Willson, 2018) and online courses (Kim & Lee, 2018). Similarly, in social sciences research, propensity score matching has been applied to assess the impact of policies and programs on various outcomes, such as labor market outcomes (Liu & Zhou, 2017).

Studies have shown that propensity score matching can improve the validity of causal inference by reducing bias and improving the balance in the distribution of covariates between treatment and control groups (Guo & Fraser, 2015; Rosenbaum, 2002). However, propensity score matching also has some limitations, including the potential for bias due to unobserved confounding variables and the difficulty of achieving balance on all covariates (Austin, 2011; Guo & Fraser, 2015). Researchers must carefully consider the assumptions and limitations of propensity score matching when applying this technique in their studies. Overall, the literature suggests that propensity score matching is a valuable tool for improving causal inference in observational studies, especially in situations where randomized experiments are not feasible or ethical (Bartlett et al., 2020; Guo & Fraser, 2015; Rosenbaum, 2002).

Data Source

The data was collected by Large Generic Community College as a part of their standard reporting practices. The data was provided in the Excel file format.

Measurement

Demographic and transcript data was provided for students who took C01 and C02 in consecutive semesters. Demographic data provided was age, ethnicity, Pell status, and gender. Transcript data was provided in the form of letter grade (A, B, C, D, F, W) and delivery method was also provided. Large Generic Community College provided all data, with data from Fall 2021 to Spring 2023.

Independent Variables

For research questions 1-3, the independent variable in each case is the delivery method, which can be either online or in-person. Each student was assigned a code based on the delivery method of their courses for both C01 and C02. These codes were IN for fully online, HY for hybrid, BL for blended, and TR for fully in-person. Based on the author's sufficient knowledge and experience with all of these delivery methods, HY was considered as online, and BL was considered as in-person for the purposes of data analysis.

Covariates

The covariates in the study consisted largely of categorical variables. The only continuous variable was age. Ethnicity was reported as White, Black, Asian, Hispanic, or other. Gender was reported as male or female. Financial aid status was reported as the student either being a Pell recipient or not.

Dependent Variables

For research question 1, the dependent variable is the demographic of the students taking the course. For research question 2, the dependent variables are the grades (as designated by letter) and the GPA (as designated by a numerical value assigned to each letter grade). For research question 3, the dependent variable is the difference between GPA from C01 to C02, indicating a rise or drop in outcome from first to second semester general chemistry.

Setting

The setting for this study is a large community college in North Carolina. This setting was chosen because of the size and availability of the data from that institution. Most community colleges offer general chemistry courses, but there may be differences in course content, course assessment, and delivery methods unique to any given university. To minimize the effects of those variables, a large college which has data for thousands of students over multiple years will provide a large dataset with internal consistency. Four-year universities are not being considered, as the institutional goals and provided experience are fundamentally different as compared to community colleges, which, as stated previously, consist largely of non-traditional students.

Population

The population was community college students enrolled in general chemistry courses General Chemistry I (C01) and General Chemistry II (C02) in consecutive semesters. Data was available for the Fall 2021 semester through the Spring 2023 semester. Large Generic Community College was able to provide data for 667 students.

Sample Size

The Sample size provided by Large Generic Community College was 667 students. There is no need to address response rate or response bias as the data was acquired as an existing dataset provided by the institution.

Sampling Criteria

Demographic data, including age, ethnicity, Pell status, and gender was provided. Transcript data was provided in the form of letter grade (A, B, C, D, F, W) and delivery method was also provided. Each subject was reported with ethnicity (White, Black, Asian, Hispanic, or other), Pell status (received aid or did not receive aid), age, gender (male or female), letter grade (A/B/C/D) and delivery method (IN/HY/BL/TR) for C01, and letter grade (A/B/C/D/F/W) and delivery method (IN/HY/BL/TR) for C01.

Data Collection

Data was obtained through Large Generic Community College, after the submission of IRB approval from NC State University and the appropriate data request forms. Large Generic Community College delivered course data in the form of excel spreadsheets, including demographic data such as age, ethnicity, Pell status, and gender. Transcript data was provided in the form of letter grade (A, B, C, D, F, W) and delivery method was also provided. disaggregated by course delivery modality.

Table 3.1

Data Collection

Area of Interest	Data Collected
Demographics	Age, Gender, Pell Status, Ethnicity
Course Grade	Transcript Data for first and second semester general chemistry
Delivery Method	Whether a course was taken online or in person

Table 3.1 describes the data that was collected

Data Source

Data sets were obtained from Large Generic Community College. Large Generic Community College provided quantitative secondary data sets, including demographic data, as well as transcript data for individual courses. The institution collects and has collected transcript data, as well as degree completion data, for students for over a decade. This data includes delivery methods. This data can be disaggregated to some extent but is limited by the information offered by participating institutions. The raw data was collected from students participating in C01 and C02 over consecutive semesters.

Pre-Data Analysis

The data was scrubbed of any student identifiers prior to analysis. The data received from Large Generic Community College delivered a data set without student names or identifiers. The identities of the students and teachers were expunged prior to being obtained from Large Generic Community College. The author did not have to do any scrubbing of the data, but the datasets were examined upon obtaining them and there were no student or teacher identifiers.

Variables were coded as categorical or continuous. Dummy variables were created for categorical data, as can be seen in Table 3.2.

Table 3.2

Covariate Variables and Coding

Variable Name	Type	Coding	
White	Dummy	1 = Yes	0 = No
Black	Dummy	1 = Yes	0 = No
Asian	Dummy	1 = Yes	0 = No
Hispanic	Dummy	1 = Yes	0 = No
Other	Dummy	1 = Yes	0 = No
Age	Continuous		
Pell Recipient	Dummy	1 = Yes	0 = No
Gender	Dummy	1 = Male	0 = Female

Table 3.2 shows how covariates were coded or recoded

The outcome variables were the grades, which were nominal, the delivery method, which was categorical, and the GPA and grade difference between C01 and C02, which were continuous. The outcome variables are summarized in Table 3.3.

Table 3.3

Outcome Variables and Coding

Variable Name	Type	Coding	
Online	Categorical	1 = Yes	0 = No
Letter Grade	Nominal	A = 4 B = 3	
		C = 2 D = 1	
		F/W = 0	
GPA	Continuous		
Grade Difference	Continuous		

Table 3.3 shows how the outcome variables were coded or recoded

Data Analysis

All data analysis was done using SPSS 29 to perform quantitative analyses. Descriptive statistics were obtained, including frequencies and percentiles. Correlations between variables were examined for statistical significance, using logistic regressions and ANOVA. The analyses were performed before and after matching via a propensity score matching methodology.

Propensity Score Matching Methodology

Propensity score matching is a quasi-experimental methodology which uses matching to control for confounding variables. Heinrich et al provide an in-depth breakdown of when and how to apply propensity score matching in *A Primer for Applying Propensity Score Matching* (2012) and Benedetto et all provide a similar breakdown in *Statistical Primer: propensity score matching and its alternatives* (2018), from which the following explanation of PSM is adapted.

The first step is to assess whether PSM is an appropriate model for the data and that using PSM will not result in biased results. Heinrich (2012) provides six criteria for model specification:

1. Any explicit criteria used in determining participation should be included in the treatment.
2. Included measures should be constant over time or deterministic with respect to time, otherwise they should be measured before participation.
3. If any particular variable has a strong influence on determining participation, it should be used in a ‘hard’ or ‘exact’ match, requiring separating subgroups based on the characteristic
4. Data for the treatment and control groups should have the same data source and collection methods.
5. Any variable with a random selection elements is good for PSM
6. Irrelevant variables which have no impact on participation should be excluded from matching.

Selecting a regression model depends on the classification of the treatment variable. Most often, these variables are dichotomous, so a logistic regression is a suitable analysis. The alternative is generally a multivariate regression, which has several drawbacks that PSM addresses, per Benedetto (2018), such as a limitation of the number of confounders and not separating the model design from the outcome analysis (p. 1113). The propensity score estimation uses the model to find the conditional probability that a subject in the control group is in the treatment group, by using a composite score generated from the given covariates. This score is calculated from the covariates with the treatment indicator used as the dependent variable, yielding a score from 0-1. These scores are used to match individuals between groups.

After a propensity score is estimated, a matching algorithm must be selected. There are several common algorithms, as detailed by Heinrich (2012, pp. 26-28). The one used in this study, nearest neighbor matching, matches individuals from the control group with the treatment group by selecting the closest propensity score, which is the most similar case. This approach can be “with replacement”, where members of the control group can match more than one member of the treatment group, or “without replacement”, where they cannot. Another approach is call radius mapping, which reduces the number of poor matches by limiting the allowable matching distance between two propensity scores, referred to as the ‘caliper’. Finally, there is Kernel and local-linear matching, which are “nonparametric matching estimators that compare the outcome of each treated person to a weighted average of the outcomes of all the untreated persons, with the highest weight being placed on those with scores closest to the treated individual.” (p. 27). These methods have low variance, but may produce poor matches. The selection of the algorithm will depend on the research context and the researcher’s discretion.

After selecting a matching algorithm, the researcher must check that the propensity score balances the covariates, testing for significance, assessing standardized bias and percent bias reduction. If the matched data is appropriate, the statistical analyses are performed on the propensity score processed data. The impact of PSM on the analysis can be seen by comparing pre- and post-match analyses.

These analyses are shown both in table format and graphically to better visualize the breakdown of the data, both pre- and post- match.

Research question 1: Demographic breakdown by course delivery method

Demographic data disaggregated by delivery method was obtained from Large Generic Community College for students who took both general chemistry courses (C101 and C102) in consecutive semesters (either Fall/Spring or Spring/Fall), excluding summers. Descriptive statistics and logistic regressions were run on the data using SPSS to assess the demographic tendencies by delivery method. The demographic data (age, ethnicity, Pell status, and gender) were used as the covariates for generating a propensity score.

Research question 2: Grade distribution by delivery method

The grading scheme used by Large Generic Community College follows the traditional letter scheme: A, B, C, D, and F, with W indicating a withdrawal. As the data set only includes students who took both semesters of general chemistry, there are no F grades in the C01 grad distribution. Large Generic Community College provided data on grade attainment for C01 and C02 sorted by delivery method. This data was evaluated with descriptive statistics and cross-tabulation, pre- and post-matching.

Research question 3: Effect of delivery method of first semester course on outcomes in second semester course

By assigning each letter grade a point value, the average change in grade from the C01 to C02 can be calculated. A grade of ‘A’ was assigned a value of 4, ‘B’ a value of 3, and so on down to ‘F’ with a value of 0. By subtracting the GPA value of C01 from C02, the grade change was shown as a whole number value, with positive numbers indicating a rise in grade from C01 to C02, and negative numbers indicating a drop in grade from C01 to C02. For example, a student who got a B in C01 and an A in C02 would have a grade difference of +1, while a student who got a B in C01 and a D in C02 would have a grade difference of -2. Descriptive statistics and ANOVA were run on the data using SPSS to assess the grade differences for each delivery method, as well as evaluating grade differences when changing delivery methods between semesters, pre- and post-matching.

Table 3.4 summarizes the research questions and the data analysis attached to each question.

Table 3.4

Research Questions and Data Analysis

Research Question	Data Analysis
1: What are the demographics that constitute in-person vs online chemistry courses of students who enrolled in consecutive semesters of general chemistry in community college?	Age was reported with mean and standard deviation Pell status, ethnicity, and gender were reported with frequency and percentile
2. Is there a statistically significant difference in grade distribution based on delivery method among students who enrolled in consecutive semesters of general chemistry in community college?	Grade distribution was reported by frequency and percent, and cross tabulated with delivery method

Table 3.4 (continued)

2a. What are the grade distributions for online general chemistry courses?	Grade distribution was reported by frequency and percent
2b. What are the grade distributions for in-person general chemistry courses chemistry courses?	Grade distribution was reported by frequency and percent
2c. Is there a statistically significant difference between a and b?	Grade distribution was cross tabulated with delivery method
3. Does the delivery method of the first semester of general chemistry impact the outcome of the second semester of general chemistry? This question involves the following:	Grade difference was calculated, and mean and standard deviation for each delivery method was calculated, as well as for changing delivery methods between semesters
3a. What are the GPA (as defined by A, B, C, D, and F) differences between first- and second- semester general chemistry for online chemistry courses?	Grade difference was calculated with mean and standard deviation
3b. What are the GPA (as defined by A, B, C, D, and F) differences between first- and second- semester general chemistry for in-person chemistry courses?	Grade difference was calculated with mean and standard deviation

Table 3.4 (continued)

3c. Is there a statistically significant difference between a and b?	ANOVA was calculated with grade difference as the dependent variable and delivery method as the factor
3d. Does changing the delivery method from first to second semester affect grade outcomes?	ANOVA was calculated with grade difference as the dependent variable and delivery method as the factor

Table 3.4 describes the research questions and the associated data analyses that were performed.

Ethical Considerations

All data was obtained from Large Generic Community College through an approved IRB request. All data was stored in password-protected folders. The data obtained from Large Generic Community College did not contain any student identifiers. Data was stored on a password-protected external hard drive, which was wiped clean after the study was completed.

Chapter Summary

This chapter discussed the data and methodology used to address the research questions. Data was provided by Large Generic Community College and included demographic and transcript data for the first two semesters of general chemistry, C01 and C02 respectively, and the delivery method. The data was analyzed using SPSS statistics software for descriptive statistics, logistic regressions, and pairwise comparisons, and ANOVA before and after a propensity-score matching screening.

Chapter 4 - Results

This study examined demographic and transcript data to determine if there were differences in academic outcomes for students at Large Generic Community College who took both semesters of general chemistry, C01 and C02, in consecutive semesters, based on the delivery method. The methodology used to examine the data was propensity-score matching (PSM), and was used to assess differences in demographics, as well as academic outcomes in terms of grade distribution, GPA, and grade change from the first to second semesters. PSM allowed for the mitigation of external factors in their effects on academic outcomes, focusing on the effect of delivery method on the academic outcomes. This chapter contains an overview of the data analysis, a description of the analysis, a summary of the outcomes and findings, and a conclusion.

Overview

PSM, as elaborated in Chapter 3, is a quasi-experimental methodology that simulates the use of a randomized sample population when a randomized sample cannot is not or cannot be utilized. PSM uses covariates to assign each individual in the sample a propensity score, then attempts to match the scores from each group that are closest to each other. This means that each sample becomes equivalent on average. In essence, “Propensity scores represent the conditional probability of a person being in one condition rather than another given a set of observed covariates used to predict a person’s condition” (Rosenbaum & Rubin, 1983, p. 4).

Data Analysis

Large Generic Community College provided samples of student data, including demographic data, transcript data for C01 and C02, and delivery method. The data was first checked for missing values and outliers and none were found. The data set contained a total of 667 students, 364 who took C01 in-person (54.6%), and 303 (45.4%) who took C01 online.

An examination of the descriptive covariates (demographic data) indicates relatively little difference in the demographic breakdown of the different delivery methods. 32% (*n* = 97) of online students received Pell grants, compared with 29.4% (*n* = 107) for in-person students, a difference of less than 3% . Similarly, online students were 38.9% (*n* = 118) male and 61.1% (*n* = 159) female while in-person students were 43.7% (*n* = 185) male and 56.3% (*n* = 205) female, a difference of less than 5%. Online students were 44.9% (*n* = 136) White, 18.8% (*n* = 57) Black, 14.9% (*n* = 45) Asian, 13.2% (*n* = 40) Hispanic, and 8.3% (*n* = 25) Other, while in-person students were 40.7% (*n* = 148) White, 16.2% (*n* = 59) Black, 14.3% (*n* = 52) Asian, 15.1% (*n* = 55) Hispanic, and 13.7% (*n* = 50) Other, differences of less than 5% in each case. Even before PSM processing, the groups are very similar in terms of demographic breakdown. Table 4.1 shows each covariate for each group prior to matching. Visualizations of the distributions of each covariate are shown in Figures 4.1 (Age), 4.2 (Pell Status), 4.3 (Gender), and 4.4 (Ethnicity).

Table 4.1

Descriptive Statistics for Online and In-Person Students in the First Semester of General Chemistry

Group Name	Online (*n* = 303)		In-Person (*n* = 364)	
Covariates	*n*	%	*n*	%
Age	303	100%	364	100%
Pell Recipient	97	32%	107	29.4%
Male	118	38.9%	159	43.7%

Table 4.1 (continued)

Female	185	61.1%	205	56.3%
White	136	44.9%	148	40.7%
Black	57	18.8%	59	16.2%
Asian	45	14.9%	52	14.3%
Hispanic	40	13.2%	55	15.1%
Other	25	8.3%	50	13.7%

Table 4.1 shows the descriptive statistics for the covariates used in the PSM assessment of the data.

Figure 4.1

Age distributions of online and in-person students pre- and post- matching

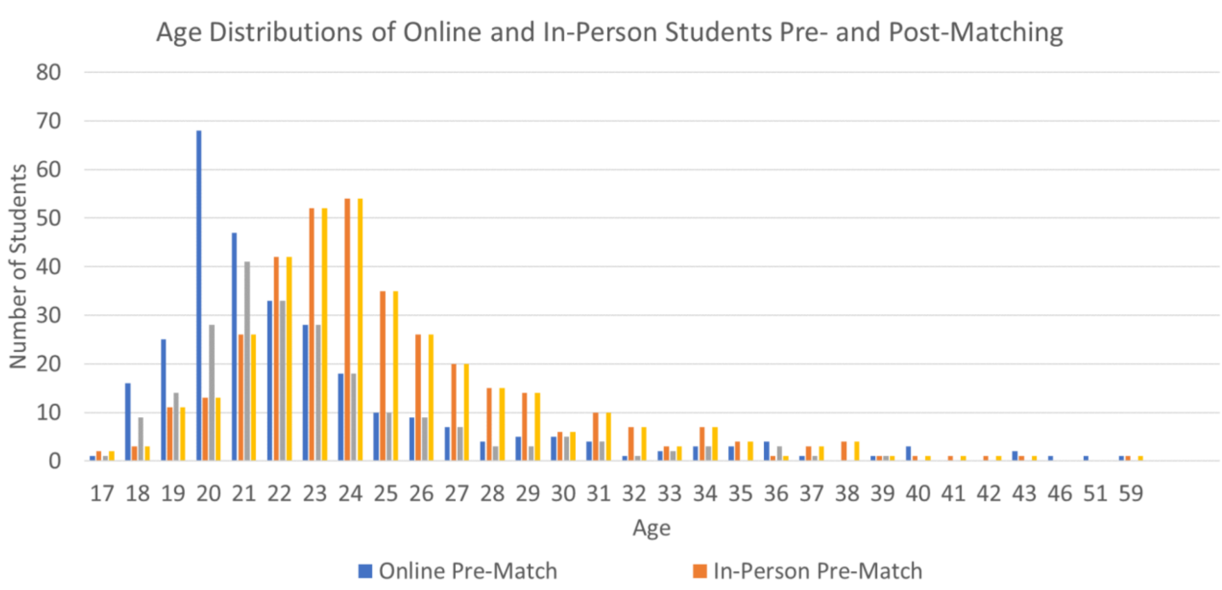


Figure 4.1 shows the age distributions for both online and in-person students, both before and after PSM was applied

Figure 4.2

Pell recipient distribution for online and in-person students pre- and post- matching

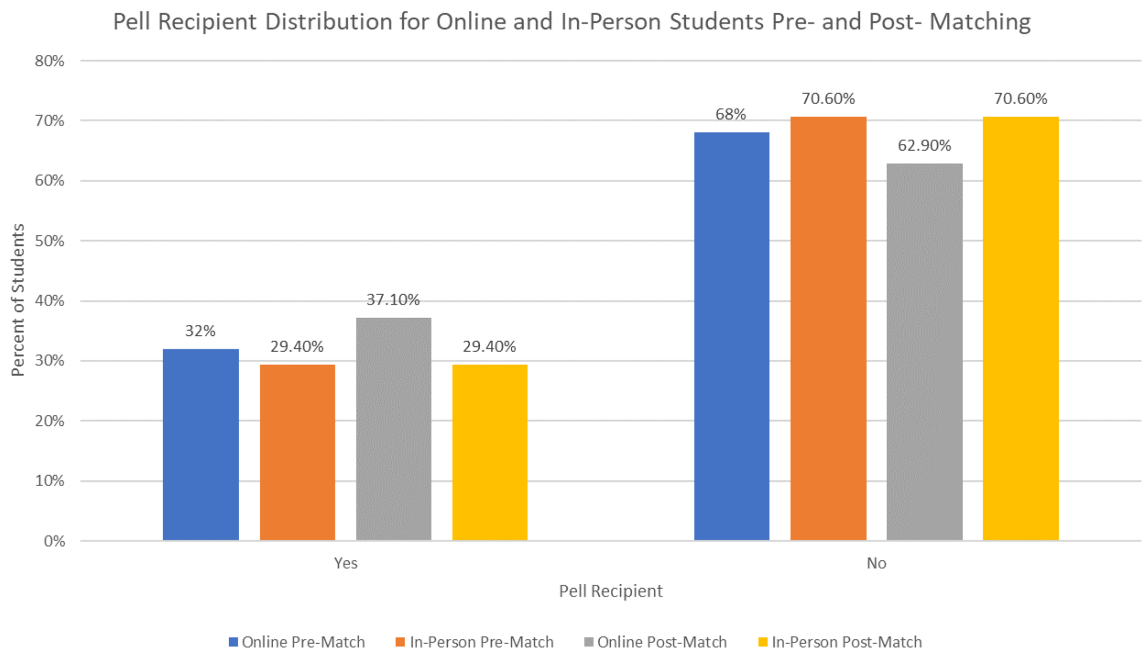


Figure 4.2 shows the Pell status distributions for both online and in-person students, both before and after PSM was applied

Figure 4.3

Gender distribution for online and in-person students pre- and post- matching

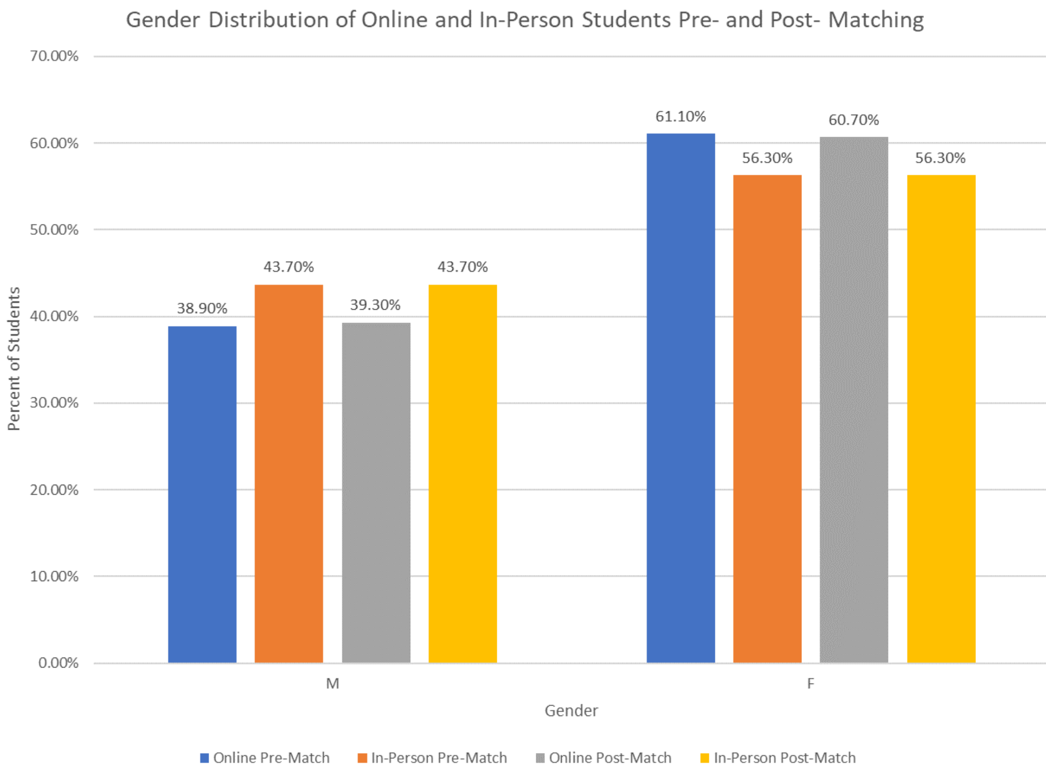


Figure 4.3 shows the gender distributions for both online and in-person students, both before and after PSM was applied

Figure 4.4

Ethnicity distribution for online and in-person students pre- and post- matching

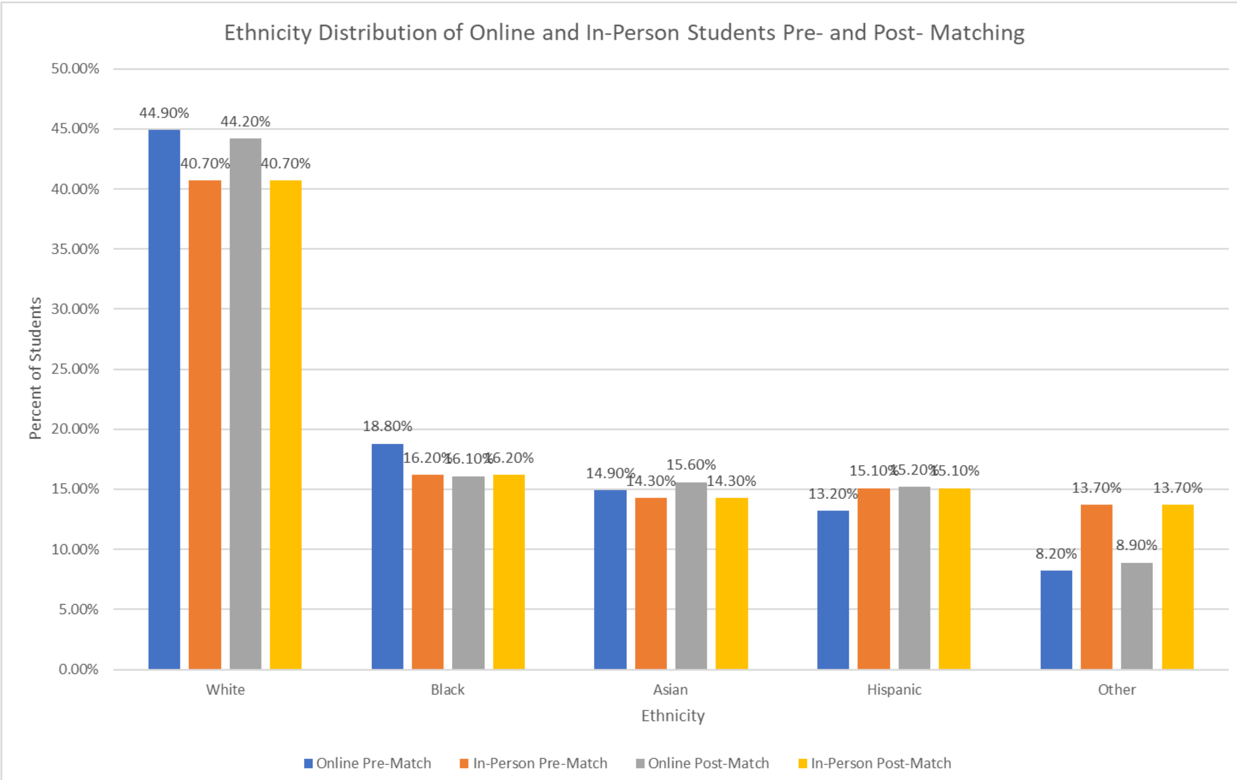


Figure 4.4 shows the ethnicity distributions for both online and in-person students, both before and after PSM was applied

The grade distributions for both C01 and C02 were assessed. Due to the population consisting exclusively of students who passed C01, there are no F or W grades for C01. The grade distributions for C01 are shown in Table 4.2. 38.9% (*n* = 118) of online students received an A, compared with 41.8% (*n* = 152) of in-person students, a difference of 1.9%. 31.7% (*n* = 96) of online students received a B, compared to 33.8% (*n* = 123) of in-person students, a difference of 2.1%. 28.7% (*n* = 87) of online students received a C, compared with 24.2% (*n* = 88) of in-person students, a difference of 4.5%. 0.7% (*n* = 2) of online students received a D, compared with 0.3% (n = 1) of in-person students, a difference of 0.4%. The C01 Grade distributions pre- and post- patching are visualized in Figure 4.5.

Table 4.2

Grade Distributions of the first semester of general chemistry

Group Name	Online (n = 303)		In-Person (n = 364)	
Grades	n	%	n	%
A	118	38.9%	152	41.8%
B	96	31.7%	123	33.8%
C	87	28.7%	88	24.2%
D	2	0.7%	1	0.3%

Table 4.2 shows the grade distributions for online and in-person students in the first semester of general chemistry, C01

Figure 4.5

C01 grade distributions for online and in-person students pre- and post- matching

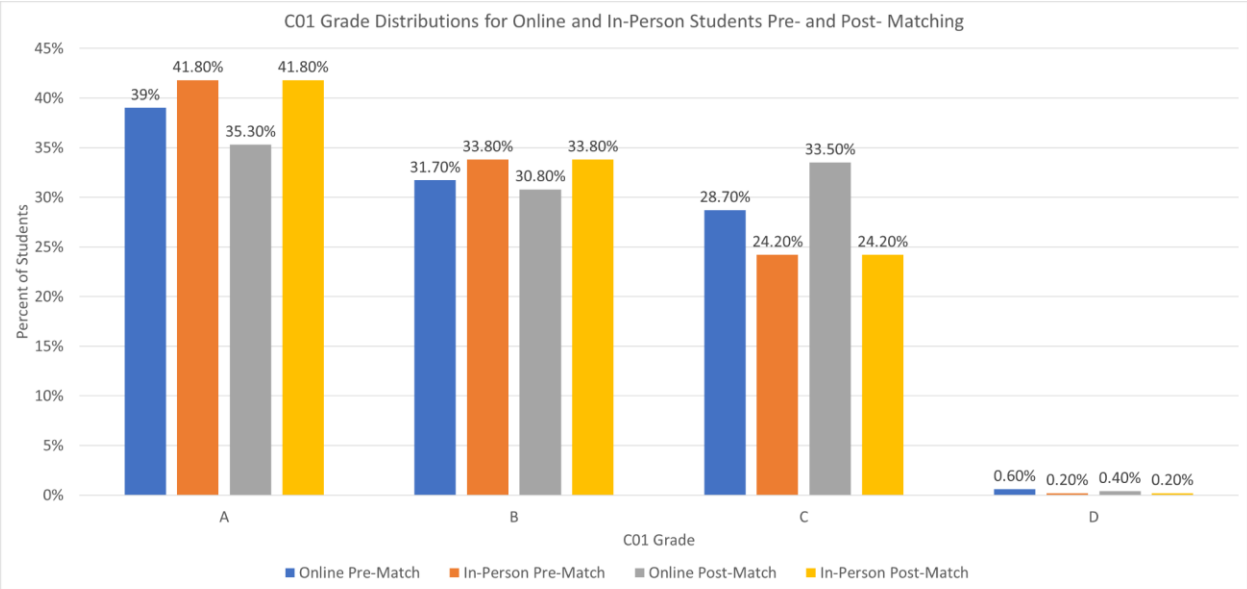


Figure 4.5 visualizes the grade distributions for online and in-person students in C01, both before and after PSM processing was applied

The grade distributions for C02 are shown in Table 4.3. 24.8% (*n* = 81) of online students received an A, compared with 26.1% (*n* = 89) of in-person students, a difference of 1.3%. 21.8% (*n* = 71) of online students received a B, compared to 25.8% (*n* = 88) of in-person students, a difference of 4%. 22.7% (*n* = 74) of online students received a C, compared with 20.8% (*n* = 71) of in-person students, a difference of 1.9%. 11.7% (*n* = 38) of online students received a D, compared with 11.4% (*n* = 39) of in-person students, a difference of 0.3%. 18.4% (*n* = 60) of online students received an F, compared to 15.8% (*n* = 54) of in-person students, a difference of 2.6%. 0.3% (*n* = 2) of online students withdraw, compared to 0% (*n* = 0) of in-person students, a difference of 0.3%. The grade distributions pre- and post- matching are visualized in Figure 4.6

Table 4.3

Grade Distributions of the second semester of general chemistry

Group Name	Online (*n* = 303)		In-Person (*n* = 364)	
Grades	*n*	%	*n*	%
A	81	24.8%	89	26.1%
B	71	21.8%	88	25.8%
C	74	22.7%	71	20.8%
D	38	11.7%	39	11.4%
F	60	18.4%	54	15.8%
W	2	0.3%	0	0.0%

Table 4.3 shows the grade distributions for online and in-person students in the first semester of general chemistry, C02

Figure 4.6

C02 grade distributions for online and in-person students pre- and post- matching

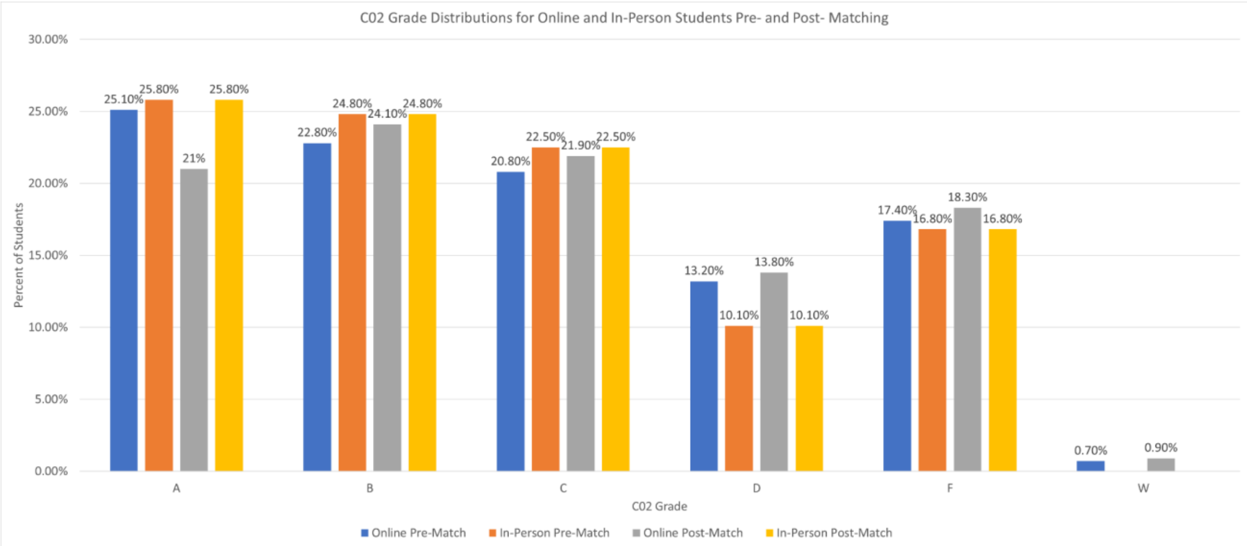


Figure 4.6 visualizes the grade distributions for online and in-person students in C02, both before and after PSM processing was applied

Propensity score matching was used in this study to mitigate the effect of any potential impact of demographics on the academic outcomes, so that the effect of the delivery method could be clearly assessed. This analysis was performed with SPSS version 29. As elucidated in Chapter 3, PSM calculates a score from 0 to 1 for each subject based on the chosen covariates. These scores are used to match subjects from each group using a nearest neighbor method, where each subject in one group is matched with the subject from another group with the closest propensity score match. If no match is found, the case is discarded, reducing the sample size. The case control match tolerance was chosen to be 0.01. After matching, 79 cases were discarded, all in the Online group, leaving the final number of cases at n = 588, with 60.5% (n = 364) in-person and 39.5% (n = 224) online. All analyses were run both pre- and post- matching to see what, if any, difference the PSM processing made to the analysis.

Logistic regression was used to assess if any of the demographic covariates had a significant correlation with whether a student took a course online or in-person. Table 4.4 shows the logistic regression of the covariates pre-matching, and table 4.5 shows the logistic regression of the covariates post-matching. Pre-matching, the model was able to correctly classify 100% of cases and accounted for 17% of the variance in the dependent variable. Wald statistics were used to test significance. None of the covariates had a significant impact on whether a student took a course online or in-person. Post-matching, the logistic regression showed some differences in ethnicity, with the online students being 15% more likely to be female, 44% more likely to be a Pell recipient, 73% more likely to be White, 53% more likely to be Black, and 69% more likely to be Asian as compared to in-person students. The only covariate that showed any significance post-matching was being a Pell recipient ($p<0.05$).

Table 4.4

Logistic regression of demographic covariates for delivery method pre-matching

Covariates	Beta	S.E.	Wald	*df*	Sig.	Exp(B)	Magnitude
Age	-0.45	0.017	7.014	1	0.008*	0.956	-4%
Gender (Male)	-.172	0.161	1.146	1	0.284	0.842	-26%
Pell Recipient	0.231	0.178	1.699	1	0.192	1.260	+26%
White	0.363	0.249	2.130	1	0.144	1.437	+44%
Black	0.329	0.285	1.329	1	0.249	1.389	+39%
Asian	0.199	0.296	0.452	1	0.501	1.220	+22%

Table 4.4 (continued)

Other	-0.258	0.161	0.621	1	0.431	0.773	-23%

Table 4.4 shows the logistic regressions for the covariates of delivery method pre-matching.

**p<0.01*

Table 4.5

Logistic regression of demographic covariates for delivery method post-matching

Covariates	Beta	S.E.	Wald	*df*	Sig.	Exp(B)	Magnitude
Age	-.065	.021	9.127	1	.003**	.937	-6%
Gender (Male)	-.164	.175	.881	1	.348	.849	-15%
Pell Recipient	.366	.184	3.928	1	.047*	1.441	+44%
White	.545	.296	3.381	1	.066	1.725	+73%
Black	.426	.340	1.570	1	.210	1.532	+53%
Asian	.524	.345	2.310	1	.129	1.689	+69%
Other	.358	.346	1.071	1	.301	1.431	+43%

Table 4.5 shows the logistic regressions for the covariates of delivery method post-matching.

**p<0.05 **p<0.01*

An analysis of variance (ANOVA) was performed on the covariates pre- and post-matching. The analysis provided means, standard deviations, F-values, significance levels, and effect size index (η^2). Age was the only covariate found to be statistically significant post-matching, F(1, 588) = 9.524, p < .01, while being a Pell recipient was almost significant, F(1. 558) = 3.729, p = .054. The effect size for age was 0.011, and the effect size for being a Pell recipient was .001, indicating the covariates had a very small effect. The pre-matching data is shown in table 4.6, and the post-matching data is shown in table 4.7.

Table 4.6

Group Mean and Standard Deviation of Covariates Before Matching

Group Name	In-Person (n = 364)		Online (n = 303)				
Covariates	M	SD	M	SD	F	*p*	η^2
Age	24.22	4.578	23.17	5.511	7.245	.007**	.011
Gender (Male)	.44	.497	.39	.488	1.527	.217	.002
Pell Recipient	.2940	.45620	.3201	.46730	.532	.466	.001
White	.4066	.49187	.4488	.49820	1.206	.273	.002
Black	.1621	.36904	.1881	.39145	.778	.378	.001
Asian	.1429	.35041	.1485	.35620	.042	.837	.000
Hispanic	.1511	.35864	.1320	.33907	.492	.483	.001
Other	.1374	.34470	.0825	.27559	5.008	.026*	.007

*Table 4.6 shows group means and standard deviations for the covariates of the delivery method pre-matching. η^2 = 0.01 indicates a small effect. * $p<0.05$ ** $p<0.01$*

Table 4.7

Group Mean and Standard Deviation of Covariates After Matching

Group Name	In-Person (n = 364)		Online (n = 224)				
Covariates	M	SD	M	SD	F	*p*	η^2
Age	24.22	4.578	23.08	3.945	9.524	.002**	.016
Gender (Male)	.44	.497	.39	.489	1.098	.295	.002
Pell Recipient	.2940	.45620	.3705	.48403	3.729	.054	.006
White	.4066	.49187	.4420	.49773	.711	.400	.001
Black	.1621	.36904	.1607	.36809	.002	.965	.000
Asian	.1429	.35041	.1563	.36391	.197	.658	.000
Hispanic	.1511	.35864	.1518	.35962	.001	.982	.000
Other	.1374	.34470	.0893	.28579	3.062	.081	.005

Table 4.7 shows group means and standard deviations for the covariates of the delivery method post-matching. $\eta^2 = 0.01$ *indicates a small effect.* * $p<0.05$ ** $p<0.01$

After matching, both the GPA statistics as well as the grade difference from C01 to C02 were analyzed with ANOVA. Each grade was assigned a point value, corresponding with the GPA that grade is worth. A grade of A was assigned four points, B three points, C two points, D one point, and W and F were assigned 0 points. The post-matching GPA statistics are shown in table 4.8. The average grade for C01 was around a B for both delivery methods (online mean of 3.01 and in-person mean of 3.17, with standard deviations of .84 and .88 respectively). The average GPA for C02 was around a C for both delivery methods (online mean of 2.13 and in-person mean of 2.34, with standard deviations of 1.42 and 1.38 respectively). Both courses showed statistically significant differences in GPA between delivery methods, in C01 (F(1,588) = 5.408, $p < .05$) and in C02 (F(1,588) = 3.909, $p < .05$), although in both cases the effect size was small (.009 for C01, .007 for C02).

Table 4.8

GPA Statistics Post-Matching

Group Name	In-Person (n = 364)		Online (n = 224)				
GPA	M	SD	M	SD	F	*p*	η^2
First Semester Post-Matching	3.1703	.80193	3.0089	.84169	5.408	.020*	.009
Second Semester Post-Matching	2.3580	1.38375	2.1288	1.41639	3.909	.049*	.007

*Table 4.8 shows the means, standard deviations, and significance of the GPA in online vs in-person courses post-matching. *p<0.05*

The grade difference from C01 to C02 was found by assigning each grade its GPA value (A = 4, B = 3, C = 2, D = 1, F/W = 0) and subtracting the C01 value from the C02 value. In this analysis, a positive number indicates an increase in grade from C01 to C02, while a negative number indicates a decrease in grade. For example, a student who received a C in C01 and a B in C02 would have a grade difference score of +1, while a student who received an A in C01 and a C in C02 would have a grade difference score of -1. This data was disaggregated on whether a student took both courses online, both courses in-person, or took C01 online and C02 in-person and vice-versa. There were no statistically significant differences between the groups, F(1,588) =

0.374, $p = 0.771$, with students on average losing almost an entire letter grade from C01 to C02. The full analysis is shown in table 4.9.

Table 4.9

Overall grade change from first to second semester based on delivery methods post-matching

Delivery Method	Results				
	n	mean	*SD*	F	sig
In-Person both semesters	301	-0.8239	1.183	0.374	0.771
Online Both Semesters	201	-0.8905	1.126		
In-Person, then Online	63	-0.9365	1.256		
Online, then in person	23	-0.6957	1.222		
Total	588	-0.8537	0.048		

Table 4.9 shows the means, standard distribution, and significance of the grade change between semesters based on delivery methods.

Summary

Chapter 4 presented the outcomes of the study, utilizing propensity-score matching to assess how online and in-person delivery methods differ in students' academic outcomes, as measured by GPA and grade difference between the first and second semesters of general chemistry. Prior to matching, there were no major differences in the demographic makeup of the online and in-person groups. After matching, the differences were virtually identical.

The logistic regression model determined that only age and Pell status were significant factors in whether a student took the courses online or in-person, and the effect size was small.

The model was able to predict with 99.5% accuracy whether a student took the courses online or in-person. There were no major differences in grade distribution between the two delivery methods, and although the post-matching analysis of the GPA averages show a statistically significant difference, the effect size is very small and as such both delivery methods had a very similar average GPA.

The grade difference between C01 and C02 was calculated to assess the impact the delivery method of C01 affects a student's academic attainment in C02. There was no notable difference found in any combination of delivery methods, whether both online or in-person, or switching delivery methods from C01 to C02. In all cases, there was a grade difference of around -.8, corresponding to an average drop of a letter grade from C01 to C02, regardless of delivery method.

The study results indicate that the delivery method of general chemistry has no significant impact on a student's academic outcomes. The study also indicated that the demographics that comprise online general chemistry courses are not markedly different from the demographics that comprise in-person chemistry courses.

Chapter 5 - Discussion and Conclusion

Chapter 5 provides conclusions, a discussion of the findings, implications for policy and practice, and recommendations for future research. The end of the chapter provides a summary of study limitations, recommendations for future research, and final conclusions.

The goal of this research was to ascertain if there were significant differences in student academic outcomes in general chemistry courses based on delivery method. Online learning has been increasing as ease of access has accelerated (Saykali, 2019), and most universities have robust online offerings (Kumar, 2017). Although this is the case, online courses in general have been found to have drawbacks in terms of student retention and completion (Gregori, 2018; Jordan, 2015). The majority of research has focused on online learning broadly, which neglects the innate differences in subject matter which make exploring the specifics of individual subjects and courses a worthwhile endeavor. General chemistry is a fairly ubiquitous course that many non-major students take (Oliver, 2021). Additionally, chemistry courses are generally considered difficult or gatekeeper courses (Sirhan, 2007; Sjöström, 2014), so research which helps student success in this area is highly desirable.

This study provided a review of the literature surrounding online education general and chemistry, as well as conceptual and theoretical frameworks for the research done. The study was conducted using propensity score matching, which method was discussed at length. This chapter details the major findings of the study, as well as a discussion of the limitations of the study and recommendations for future research.

Discussion

The research questions posed in this study focused on the potential differences in population and academic outcomes between students taking general chemistry online versus those taking the course in-person. The findings of the research elucidated the impact that the delivery method of a course has on a student's success within the course and in future courses, as measured by GPA and grade difference between the two semesters.

Research Question 1

Research question 1 examined the demographics of students taking general chemistry courses online as compared to students taking general chemistry courses in-person. All the students in the study took both semesters of general chemistry in consecutive semesters, between Fall 2021 and Spring 2023. Records for 667 students were provided, with 303 (45.4%) taking the courses online and 364 (54.6%) taking the courses in-person. 32% of the online students and 29.4% of the in-person students were Pell recipients. The online courses were 38.9% male and 61.1% female while the in-person courses were 43.7% male and 56.3% female. The most represented ethnicities were White at 44.9% online and 40.7% in-person, followed by Black at 18.8% online and 14.3% in-person. The demographics were each within 5% of each other across groups, indicating that there were no major differences between the groups. Logistic regression showed that age and Pell status were significant factors, but their effect size was very small.

Research Question 2

Research question 2 assessed the differences in academic outcomes between the online and in-person groups by looking at grade distributions and GPA averages for the general chemistry courses. The grade distributions were very similar for both semesters, with 38.9% of online students receiving an A in the first semester, compared to 41.8% of in-person students. In general, online students were a few percentage points behind their in-person counterparts, but the differences were slim, with online students receiving 2.9% fewer A grades, 2.1% fewer B grades, 4.3% more C grades, and 0.4% more D grades. The second semester grades followed similar patterns, with online students performing just below their counterparts in in-person courses, with online students receiving 1.3% fewer A grades, 4% fewer B grades, 1.9% more C grades, 2.6% more D grades, 2.6% more F grades, and 0.3% more withdrawals. The grade distributions were very similar, indicating that there is no apparent difference in outcome between delivery methods.

Each grade was assigned a standard GPA point value (A = 4, B = 3, C = 2, D = 1, F/W = 0) and the GPA values were analyzed. After matching, the average first semester GPA was 3.00±.84 for the online group and 3.17±.80 for the in-person group, while the average second semester GPA was 2.12±1.41 for the online group and 2.36±1.38 for the in-person group. Both these statistics were significant, but the effect size was very small, indicating that in-person courses lead to very slightly improved outcomes.

Research Question 3

Research question 3 assesses whether the delivery method of the first semester impacted the academic outcome of the second semester. This was done by assigning each grade a GPA value as discussed previously, then subtracting the first semester GPA from the second semester GPA. In this manner, a positive grade difference indicates a rise in grade from first to second semester, and a negative grade difference indicates a drop in grade. This data was then analyzed for every combination of delivery methods across the two semesters: both online, both in-person, first semester online and second in-person, and first semester in-person and second online. The results showed no meaningful difference between delivery methods, with all students dropping around a full letter grade from first to second semester (both online, -.89 difference; both in-person, -.82 difference; in-person, then online, -.94 difference; online, then in-person, -.70 difference; overall, -.85 difference). This indicates that the delivery method of an earlier chemistry course does not impact a student's academic outcome in future chemistry courses, although on average the outcomes drop in more advanced courses.

Implications for Policy

The results of this research may or may not have impacts on policy, depending on the context of policies surrounding online courses. To the author's knowledge, there are no institutions which weight courses differently in a transcript based on delivery method, to this author's knowledge the delivery method is not shown in official transcripts. However, the results of the study, although there ought to be more research done on larger sample sizes to get a clearer picture of the impact of delivery methods, should give institutions confidence that online courses are just as meritorious as in-person courses, and deserve similar backing and resources, as they are an appealing alternative to many students.

Implications for Practice

While the implications for practice are subject to context, in general, from the data it seems reasonable to continue offering g general chemistry courses online and perhaps expanding said offerings, since general chemistry is such a highly attended course.It remains to be seen if the success of online education in general chemistry courses carries through to more advanced courses such as organic chemistry.

Recommendations for future research

This study was just a glimpse into the topic of online chemistry education, much more research is needed in this area to further explore the effects of delivery methods on student success on a per-course basis. The results of this study point to a conclusion that there is no significant difference in academic outcomes for general chemistry students between online and in-person delivery methods. Additionally, the delivery method of earlier courses seems to have no or negligible impacts on student academic outcomes in future chemistry courses. The data did show a very marginal decrease in academic outcomes for online students. The author strongly recommends this research be continued on larger sample sizes across more courses. The literature review showed that research into the impact on delivery methods on academic outcomes were often extremely broad in scope, not taking into account the differences between subject areas. There were few, if any, studies that examined general chemistry specifically, and the author was unable to find any literature examining the impact of delivery methods in earlier courses on outcomes in more advanced courses.

Recommendation 1

The first and foremost recommendation is to expand the size and scope of the data set analyzed. Propensity-score matching is a robust tool, and becomes more robust as more covariates and larger sample sizes are available. Many schools already collect this data, so it is simply a matter of collecting and analyzing this. Conducting a much larger scale examination, with multiple schools and institutions involved, would help assess general chemistry broadly, as it is a ubiquitous course, as mentioned previously. Data ought to be examined and parsed at national, state, and institution levels, as this would aid institutions in assessing how their own chemistry programs align with national standards. Additionally, providing a covariate of the instructor for the class would help mitigate any variation due to individual teachers.

Recommendation 2

As just mentioned, the literature did not provide many examples of course-specific analyses of delivery methods. It would be interesting to examine more courses, both within similar subject areas (such as physics and biology), and in different subject areas (such as English or writing courses). It would be especially interesting to see if similar patterns emerge between courses that contain laboratory components and those that do not. As different subject matters will require different applications of knowledge and skills which translate to different classroom methodologies, a meta-analysis of delivery methods between subject areas would provide fascinating data on how to optimize online education in different subject areas.

Recommendation 3

This study examined how delivery methods in earlier classes may affect outcomes in more advanced courses, which to this author's knowledge has not been done before. Even at the institution which provided the data, there were more advanced chemistry courses offered. It would be fascinating to assess if the delivery method of general chemistry impacts student outcomes in organic chemistry, a notoriously difficult course. The lack of this data was one limitation of this study and the results of examining four chemistry courses would provide a stronger conclusion than merely looking at two.

Recommendation 4

All of the data provided was obtained post-COVID, an event that radically altered the education landscape. While the full effects of COVID will likely not be known for years to come, it is certain that the viability of online education was pushed to the forefront. An analysis of online versus in-person academic performance pre-, during, and post- COVID would provide a fascinating insight into how online education has changed over the years, and how the disruptive impact of a global pandemic affected the educational landscape.

Limitations

The biggest limitation to the study was the number of students in the data set. Large Generic Community College was only able to provide a small subset of the students who took general chemistry, further pared down to only students who took both semesters consecutively. This means the data set inherently eliminated any students who did not pass the first semester or who chose to forgo the second semester. The author was therefore unable to assess whether the grade distributions changed between semesters, since the first semester data was lacking most D grades and all F/W grades.

The second limitation was the lack of data for students who took organic chemistry. Large Generic Community College offers two semesters of organic chemistry and while these courses do not have as high enrollment as the general chemistry courses, having the data for those students who did take organic chemistry would have clarified the potential impact delivery method in earlier courses has on later courses.

These limitations were largely due to the relative dearth of data that was available. Any variations due to non-random sampling was mitigated by the use of propensity-score matching.

Conclusion

The results of this study were quite surprising to the author. The literature seemed to indicate that, in general, online courses tended to have worse outcomes for students. The author's intuition and experience with general chemistry would seem to back up those assumptions, as general chemistry is a difficult course which frequently requires students to seek direct help from the instructor, which is generally more difficult in online courses. The fact that the demographics of the online and in-person groups were practically identical was unexpected, as the students choose online and in-person courses for different reasons, so some difference in demographics was expected but not observed. The outcomes in terms of GPA and grade distribution were very similar, which could possibly be attributed to online instructors being more lenient with grading, if there was in fact a difference in learning, but this is countered by the fact that the grade difference was almost identical, regardless of the delivery method. Even changing delivery methods had no impact on grade outcomes.

Ultimately, these results should be seen as very positive. As an instructor, I intuitively feel that online courses, especially in the hard sciences, do not equip students as well as in-person courses. I am glad to have my assumptions overturned and to see that online chemistry education is a viable strategy.

References

Ahn, J. N., Hu, D., & Vega, M. (2020). “Do as I do, not as I say”: Using social learning theory to unpack the impact of role models on students' outcomes in education. *Social and Personality Psychology Compass*, *14*(2), e12517.

Akçay, H., Durmaz, A., Tüysüz, C., & Feyz, B. (2006). EFFECTS OF COMPUTER BASED LEARNING ON STUDENTS’ ATTITUDES AND ACHIEVEMENTS TOWARDS ANALYTICAL CHEMISTRY. *The Turkish Online Journal of Educational Technology*, *5*(1), 5.

Alam, A. (2022). Mapping a sustainable future through conceptualization of transformative learning framework, education for sustainable development, critical reflection, and responsible citizenship: an exploration of pedagogies for twenty-first century learning. *ECS Transactions*, *107*(1), 9827.

Allen, I. E. (2010). *Class Differences: Online Education in the United States, 2010*. 30.

Allen, I. E., Seaman, J., Sloan Consortium, Babson Survey Research Group, & Pearson Foundation. (2013). *Changing Course Ten Years of Tracking Online Education in the United States*. Distributed by ERIC Clearinghouse.

Andragogy Homepage for Adult Education Specialists. (n.d.). Retrieved February 27, 2022, from http://www.andragogy.net/

Antunes, M., Pacheco, M. A. R., & Giovanela, M. (2012). Design and Implementation of an Educational Game for Teaching Chemistry in Higher Education. *Journal of Chemical Education*, *89*(4), 517–521. https://doi.org/10.1021/ed2003077\

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behavioral Research, 46(3), 399-424. doi: 10.1080/00273171.2011.568786

Bambara, C. S., Harbour, C. P., Davies, T. G., & Athey, S. (2009). Delicate engagement: The lived experience of community college students enrolled in high-risk online courses. *Community College Review*, *36*(3), 219-238.

Bandura, A., & Walters, R. H. (1977) *Social learning theory* (Vol. 1). Prentice Hall: Englewood cliffs

Barak, M. (2004). *Enhancing undergraduate students' chemistry understanding through project-based learning in an IT environment*. https://doi.org/10.1002/sce.20027

Bartlett, J. W., Harel, O., Zhou, X.-H., & Kahler, K. H. (2020). Beyond hypothesis testing: A primer on propensity score analysis in pharmacoepidemiology. Current Epidemiology Reports, 7(4), 338-346. doi: 10.1007/s40471-020-00243-1

Belfield, C. R., & Bailey, T. (2011). The benefits of attending community college: A review of the evidence. *Community College Review*, *39*(1), 46-68.

Benedetto, U., Head, S. J., Angelini, G. D., & Blackstone, E. H. (2018). Statistical primer: propensity score matching and its alternatives. *European Journal of Cardio-Thoracic Surgery*, *53*(6), 1112-1117.

Bergman, M., Gross, J. P., Berry, M., & Shuck, B. (2014). If life happened but a degree didn't: Examining factors that impact adult student persistence. *The Journal of Continuing Higher Education*, *62*(2), 90-101.

Bettinger, E., & Loeb, S. (2017). *Promises and pitfalls of online education*. 4.

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), 2855–2875. https://doi.org/10.1257/aer.20151193

Bird, K. A., Castleman, B. L., & Lohner, G. (2022). Negative Impacts From the Shift to Online Learning During the COVID-19 Crisis: Evidence From a Statewide Community College System. *AERA Open*, *8*, 233285842210812. https://doi.org/10.1177/23328584221081220

Bowen, W. G. (2015). Higher education in the digital age. In *Higher Education in the Digital Age*. Princeton University Press.

Bragg, D. D. (2001). Opportunities and challenges for the new vocationalism in American community colleges. *New directions for community colleges*, *2001*(115), 5-15.

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, *27*, 1–13. https://doi.org/10.1016/j.iheduc.2015.04.007

Brown, J. S. (2001). *Learning in the Digital Age*. 27.

Brown, John Seely, D., Paul. (1996). Universities in the digital age. *The Magazine of Higher Learning*, *28*(4), 11–19. https://doi-org.prox.lib.ncsu.edu/10.1080/00091383.1996.9937757

Bucat, R. (2004). PEDAGOGICAL CONTENT KNOWLEDGE AS A WAY FORWARD: APPLIED RESEARCH IN CHEMISTRY EDUCATION. *Chem. Educ. Res. Pract.*, *5*(3), 215–228. https://doi.org/10.1039/B4RP90025A

Bulte, A. M. W., Westbroek, H. B., de Jong, O., & Pilot, A. (2006). A Research Approach to Designing Chemistry Education using Authentic Practices as Contexts. *International Journal of Science Education*, *28*(9), 1063–1086. https://doi.org/10.1080/09500690600702520

Carr, S. (2000). As Distance Education Comes of Age, the Challenge Is Keeping the Students. *Chronicle of Higher Education*, *46*(23).

Carter, C. S., & Brickhouse, N. W. (1989). What makes chemistry difficult? Alternate perceptions. *Journal of Chemical Education*, *66*(3), 223. https://doi.org/10.1021/ed066p223

Chametzky, B. (2014). Andragogy and Engagement in Online Learning: Tenets and Solutions. *Creative Education*, *05*(10), 813–821. https://doi.org/10.4236/ce.2014.510095

Chittleborough, G. D., Treagust, D. F., & Mocerino, M. (2007). Achieving Greater Feedback and Flexibility Using Online Pre-Laboratory Exercises with Non-Major Chemistry Students. *Journal of Chemical Education*, *84*(5), 884. https://doi.org/10.1021/ed084p884

Coley, N. R. (1973). Prediction of success in general chemistry in a community college. *Journal of Chemical Education*, *50*(9), 613. https://doi.org/10.1021/ed050p613

Cooper, M. M., & Stowe, R. L. (2018). Chemistry Education Research—From Personal Empiricism to Evidence, Theory, and Informed Practice. *Chemical Reviews*, *118*(12), 6053–6087. https://doi.org/10.1021/acs.chemrev.8b00020

Coros, J. D., & Madrigal, D. V. (2021). Self-directed learning, self-efficacy in learning, and academic motivation of public senior high school students. *Asian Journal of Education and Social Studies*, *21*(2), 19-34.

Cox, R. D. (2005). *Online Education as Institutional Myth: Rituals and Realities at Community Colleges*.

Crittenden, W. F., Biel, I. K., & Lovely III, W. A. (2019). Embracing digitalization: Student learning and new technologies. *Journal of marketing education*, *41*(1), 5-14.

Dalgarno, B., Bishop, A. G., Adlong, W., & Bedgood, D. R. (2009). Effectiveness of a Virtual Laboratory as a preparatory resource for Distance Education chemistry students. *Computers & Education*, *53*(3), 853–865. https://doi.org/10.1016/j.compedu.2009.05.005

DECELLE, G. (2016). Andragogy: A Fundamental Principle of Online Education for Nursing. *Journal of Best Practices in Health Professions Diversity*, *9*(2), 1263–1273.

DeTure, M. (2004). Cognitive Style and Self-Efficacy: Predicting Student Success in Online Distance Education. *American Journal of Distance Education*, *18*(1), 21–38. https://doi.org/10.1207/s15389286ajde1801_3

Eichler, J. F., & Peeples, J. (2013). Online Homework Put to the Test: A Report on the Impact of Two Online Learning Systems on Student Performance in General Chemistry. *Journal of Chemical Education*, *90*(9), 1137–1143. https://doi.org/10.1021/ed3006264

Eilks, I., & Byers, B. (2010). The need for innovative methods of teaching and learning chemistry in higher education – reflections from a project of the European Chemistry

Thematic Network. *Chem. Educ. Res. Pract.*, *11*(4), 233–240. https://doi.org/10.1039/C0RP90004D

Ferreira, C. C. (2020). Experiential learning theory and hybrid entrepreneurship: factors influencing the transition to full-time entrepreneurship. *International Journal of Entrepreneurial Behavior & Research*, *26*(8), 1845-1863.

Gabrielle, D. (2003). *The Effects of Technology-Mediated Instructional Strategies on Motivation, Performance, and Self-Directed Learning*. 2568–2575. http://www.learntechlib.org/primary/p/14267/

Garcia-Martinez, J. (2010). *Chemistry 2.0 Creating Online Communities*. 4.

Garrison, R. (2000). Theoretical Challenges for Distance Education in the 21st Century: A Shift from Structural to Transactional Issues. *The International Review of Research in Open and Distributed Learning*, *1*(1). https://doi.org/10.19173/irrodl.v1i1.2

Geng, S., Law, K. M., & Niu, B. (2019). Investigating self-directed learning and technology readiness in blending learning environment. *International Journal of Educational Technology in Higher Education*, *16*(1), 1-22.

Goldrick-Rab, S. (2010). Challenges and opportunities for improving community college student success. *Review of educational research*, *80*(3), 437-469.

Gregori, P., Martínez, V., & Moyano-Fernández, J. J. (2018). Basic actions to reduce dropout rates in distance learning. *Evaluation and program planning*, *66*, 48-52.

Grove, N., & Bretz, S. L. (2007). CHEMX: An Instrument To Assess Students' Cognitive Expectations for Learning Chemistry. *Journal of Chemical Education*, *84*(9), 1524. https://doi.org/10.1021/ed084p1524

Guo, S., & Fraser, M. W. (2015). Propensity score analysis: Statistical methods and applications (2nd ed.). Los Angeles, CA: SAGE Publications.

Harasim, L. (2000). Shift happens: Online education as a new paradigm in learning. *The Internet and Higher Education*, *3*(1–2), 41–61. https://doi.org/10.1016/S1096-7516(00)00032-4

Hartree, A. (2006). Malcolm Knowles' Theory of Andragogy: A Critique. *International Journal of Lifelong Education*. https://doi.org/10.1080/0260137840030304

He, Y., Swenson, S., & Lents, N. (2012). Online Video Tutorials Increase Learning of Difficult Concepts in an Undergraduate Analytical Chemistry Course. *Journal of Chemical Education*, *89*(9), 1128–1132. https://doi.org/10.1021/ed200685p

Heinrich, C., Maffioli, A., & Vazquez, G. (2010). A primer for applying propensity-score matching.

Hobson, T. D., & Puruhito, K. K. (2018). Going the distance: Online course performance and motivation of distance learning students. *Online Learning*, *22*(4), 129-140.

Ives, K. S. (2019). COMMUNITY COLLEGES AND DISTANCE LEARNING. *Online Learning*, *10*(3). https://doi.org/10.24059/olj.v10i3.1757

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

Johnstone, A. H. (2000). TEACHING OF CHEMISTRY - LOGICAL OR PSYCHOLOGICAL? *Chem. Educ. Res. Pract.*, *1*(1), 9–15. https://doi.org/10.1039/A9RP90001B

Jordan, K. (2015). Massive open online course completion rates revisited: Assessment, length and attrition. *The International Review of Research in Open and Distributed Learning*, *16*(3). https://doi.org/10.19173/irrodl.v16i3.2112

JRCET vol 4 no 2. (n.d.). Retrieved February 23, 2021, from http://rcetj.org/index.php/rcetj/article/view/22/28

Kauffman, H. (2015). A review of predictive factors of student success in and satisfaction with online learning. *Research in Learning Technology, 23*. https://doi.org/10.3402/rlt.v23.26507

Kennet, L. N. (2015). *Instructor Presence Helps Bridge the Gap between Online and On-Campus Learning*. https://files.eric.ed.gov/fulltext/EJ1095942.pdf

Kim, Y., & Lee, J. (2018). Propensity score matching analysis of online versus face-to-face high school calculus course. Computers & Education, 118, 1-15. doi: 10.1016/j.compedu.2017.11.009

King, D. (2012). New perspectives on context-based chemistry education: Using a dialectical sociocultural approach to view teaching and learning. *Studies in Science Education*, *48*(1), 51–87. https://doi.org/10.1080/03057267.2012.655037

Knowles, M. S. (1978). Andragogy: Adult Learning Theory in Perspective. *Community College Review*, *5*(3), 9–20. https://doi.org/10.1177/009155217800500302

Knowles, M. S. (1984). Theory of andragogy.

Kolb, David A. *Experiential learning: Experience as the source of learning and development*. FT press, 2014.

Kolack, K., Hemraj-Benny, T., & Chauhan, M. (2020). Community College Chemistry Instruction and Research in the Time of COVID-19. *Journal of Chemical Education*, *97*(9), 2889–2894. https://doi.org/10.1021/acs.jchemed.0c00700

Kumar, A., Kumar, P., Palvia, S. C. J., & Verma, S. (2017). Online education worldwide: Current status and emerging trends. *Journal of Information Technology Case and Application Research*, *19*(1), 3-9.

Ladwig, A. (2021). *Online Learning in Graduate Health Programs: Andragogy, Lecture Preference, and the Effectiveness of Synchronous and Asynchronous Participation* (Doctoral dissertation, University of South Dakota).

Lane, F., To, Y., Shelley, K., & Henson, R. (2012). An illustrative example of propensity score matching with education research. *Career and Technical Education Research*, *37*(3), 187-212.

Larreamendy-Joerns, J., & Leinhardt, G. (2006). Going the Distance With Online Education. *Review of Educational Research*, *76*(4), 567–605. https://doi.org/10.3102/00346543076004567

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press.

Leontyev, A., & Baranov, D. (2013). Massive Open Online Courses in Chemistry: A Comparative Overview of Platforms and Features. *Journal of Chemical Education*, *90*(11), 1533–1539. https://doi.org/10.1021/ed400283x

Levy, Y. (2007). *Comparing dropouts and persistence in e-learning courses*. 20.

Lewin, L. O., McManamon, A., Stein, M. T., & Chen, D. T. (2019). Minding the form that transforms: using Kegan's model of adult development to understand personal and professional identity formation in medicine. *Academic medicine*, *94*(9), 1299-1304.

Li, A., & Hamlin, D. (2019). Is daily parental help with homework helpful? Reanalyzing national data using a propensity score–based approach. *Sociology of Education*, *92*(4), 367-385.

Littlejohn, A., & Hood, N. (2018). *Reconceptualising Learning in the Digital Age*. Springer Singapore. https://doi.org/10.1007/978-981-10-8893-3

Liu, S., & Zhou, X.-H. (2017). Propensity score matching and statistical inference. International Statistical Review, 85(1), 98-118. doi: 10.1111/insr.12154

Lloyd, P. M., & Eckhardt, R. A. (2010). *Strategies for Improving Retention of Community College Students In the Sciences*. *19*(1).

Longmore, D. J. (2021). TRANSFORMATIVE LISTENING ACROSS GLOBAL CONTEXTS: SITUATED LEARNING THEORY AND THE USE OF A LISTENING PROTOCOL TO CREATE SAFE SPACES FOR LISTENING AND CONNECTION. *Actes de*, 22.

Mahaffy, P. (2004). THE FUTURE SHAPE OF CHEMISTRY EDUCATION. *Chem. Educ. Res. Pract.*, *5*(3), 229–245. https://doi.org/10.1039/B4RP90026J

Mahaffy, P. G., Krief, A., Hopf, H., Mehta, G., & Matlin, S. A. (2018). *Reorienting chemistry education through systems thinking*. 3.

McGee, P. (2013). Supporting Academic Honesty in Online Courses. *Journal of Educators Online*, *10*(1), 1–31.

McKinney, L., Novak, H., Hagedorn, L. S., & Luna-Torres, M. (2019). Giving up on a course: An analysis of course dropping behaviors among community college students. *Research in Higher Education*, *60*, 184-202.

Means, B., Toyama, Y., Murphy, R., & Baki, M. (2013). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers college record*, *115*(3), 1-47.

Mews, J. (2020). Leading through Andragogy. *College and University*, *95*(1), 65–68.

Meyer, L. S., Panee, D., Schmidt, S., & Nozawa, F. (2003). Using Demonstrations to Promote Student Comprehension in Chemistry. *Journal of Chemical Education*, *80*(4), 431. https://doi.org/10.1021/ed080p431

Moloney, J. F., & Oakley, B. (2019). SCALING ONLINE EDUCATION: INCREASING ACCESS TO HIGHER EDUCATION. *Online Learning*, *14*(1). https://doi.org/10.24059/olj.v14i1.1639

Moore, E. B., Chamberlain, J. M., Parson, R., & Perkins, K. K. (2014). PhET Interactive Simulations: Transformative Tools for Teaching Chemistry. *Journal of Chemical Education*, *91*(8), 1191–1197. https://doi.org/10.1021/ed4005084

Moseley, C., Summerford, H., Paschke, M., Parks, C., & Utley, J. (2020). Road to collaboration: Experiential learning theory as a framework for environmental education program development. *Applied Environmental Education & Communication*, *19*(3), 238-258.

Mukhalalati, B. A., & Taylor, A. (2019). Adult Learning Theories in Context: A Quick Guide for Healthcare Professional Educators. *Journal of Medical Education and*

Curricular Development, *6*, 238212051984033. https://doi.org/10.1177/2382120519840332

Mulford, D. R., & Robinson, W. R. (2002). An Inventory for Alternate Conceptions among First-Semester General Chemistry Students. *Journal of Chemical Education*, *79*(6), 739. https://doi.org/10.1021/ed079p739

Murphy, C. A., & Stewart, J. C. (2015). *The Impact of Online or F2F Lecture Choice on Student Achievement and Engagement in a Large Lecture-Based Science Course: Closing the Gap*. *19*(3), 20.

Ogot, M., Elliott, G., & Glumac, N. (2003). An Assessment of In-Person and Remotely Operated Laboratories. *Journal of Engineering Education*, *92*(1), 57–64. https://doi.org/10.1002/j.2168-9830.2003.tb00738.x

Oseguera, L., & Rhee, B. S. (2009). The Influence of Institutional Retention Climates on Student Persistence to Degree Completion: A Multilevel Approach. *Research in Higher Education*, *50*(6), 546–569. https://doi.org/10.1007/s11162-009-9134-y

Parker, L. L., & Loudon, G. M. (2013). Case Study Using Online Homework in Undergraduate Organic Chemistry: Results and Student Attitudes. *Journal of Chemical Education*, *90*(1), 37–44. https://doi.org/10.1021/ed300270t

Paulsen, J., & McCormick, A. C. (2020). Reassessing disparities in online learner student engagement in higher education. *Educational Researcher*, *49*(1), 20-29.

Perera, V., Mead, C., Buxner, S., Lopatto, D., Horodyskyj, L., Semken, S., & Anbar, A. D. (2017). Students in Fully Online Programs Report More Positive Attitudes toward Science Than Students in Traditional, In-Person Programs. *CBE—Life Sciences Education*, *16*(4), ar60. https://doi.org/10.1187/cbe.16-11-0316

Philibert, N., Allen, J., & Elleven, R. (2008). Nontraditional Students in Community Colleges and the Model of College Outcomes for Adults. *Community College Journal of Research and Practice*, *32*(8), 582–596. https://doi.org/10.1080/10668920600859913

Picciano, A. G. (2019). BEYOND STUDENT PERCEPTIONS: ISSUES OF INTERACTION, PRESENCE, AND PERFORMANCE IN AN ONLINE COURSE. *Online Learning*, *6*(1). https://doi.org/10.24059/olj.v6i1.1870

Powers, R. A., Cochran, J. K., Maskaly, J., & Sellers, C. S. (2020). Social learning theory, gender, and intimate partner violent victimization: A structural equations approach. *Journal of interpersonal violence*, *35*(17-18), 3554-3580.

Prendergast, G. A. (2003). Keeping online student dropout numbers low. *Global-Ed. com*.

Radovan, M. (2019). Should I stay, or should I go? Revisiting student retention models in distance education. *Turkish Online Journal of Distance Education*, *20*(3), 29-40.

Rastegarpour, H., & Marashi, P. (2012). The effect of card games and computer games on learning of chemistry concepts. *Procedia - Social and Behavioral Sciences*, *31*, 597–601. https://doi.org/10.1016/j.sbspro.2011.12.111

Redmond, P., Abawi, L., Brown, A., Henderson, R., & Heffernan, A. (2018). An online engagement framework for higher education. *Online learning*, *22*(1), 183-204.

Reid, N. (2000). THE PRESENTATION OF CHEMISTRY LOGICALLY DRIVEN OR APPLICATIONS-LED? *Chem. Educ. Res. Pract.*, *1*(3), 381–392. https://doi.org/10.1039/B0RP90018D

Reid, N., & Shah, I. (2007). The role of laboratory work in university chemistry. *Chem. Educ. Res. Pract.*, *8*(2), 172–185. https://doi.org/10.1039/B5RP90026C

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.

Rosenbaum, P. R. (2002). Observational studies (2nd ed.). New York, NY: Springer.

Rossman, M. H. (2000). Andragogy and Distance Education: Together in the New Millennium. *New Horizons in Adult Education and Human Resource Development*, *14*(1), 4–11. https://doi.org/10.1002/nha3.10105

Sarvary, M. A., Castelli, F. R., & Asgari, M. (2022). Undergraduates' experiences with online and in-person courses provide opportunities for improving student-centered biology laboratory instruction. *Journal of Microbiology & Biology Education*, *23*(1), e00289-21.

Sato, T., Haegele, J. A., & Foot, R. (2017). Developing Online Graduate Coursework in Adapted Physical Education Utilizing Andragogy Theory. *Quest*, *69*(4), 453–466. https://doi.org/10.1080/00336297.2017.1284679

Sawatsky, A. P., Ratelle, J. T., Bonnes, S. L., Egginton, J. S., & Beckman, T. J. (2017). A model of self-directed learning in internal medicine residency: a qualitative study using grounded theory. *BMC medical education*, *17*(1), 1-9.

Saykili, A. (2019). Higher education in the digital age: The impact of digital connective technologies. *Journal of Educational Technology and Online Learning*, *2*(1), 1-15.

Scheunemann, A., Schnettler, T., Bobe, J., Fries, S., & Grunschel, C. (2022). A longitudinal analysis of the reciprocal relationship between academic procrastination, study satisfaction, and dropout intentions in higher education. *European Journal of Psychology of Education*, *37*(4), 1141-1164.

Seery, M. K. (2015). Flipped learning in higher education chemistry: Emerging trends and potential directions. *Chemistry Education Research and Practice*, *16*(4), 758–768. https://doi.org/10.1039/C5RP00136F

Singhal, S., Bagga, S., Goyal, P., & Saxena, V. (2012). Augmented Chemistry: Interactive Education System. *International Journal of Computer Applications*, *49*(15), 1–5. https://doi.org/10.5120/7700-1041

Sirhan, G. (2007). *Learning Difficulties in Chemistry: An Overview*. 19.

Sjöström, J., & Talanquer, V. (2014). Humanizing Chemistry Education: From Simple Contextualization to Multifaceted Problematization. *Journal of Chemical Education*, *91*(8), 1125–1131. https://doi.org/10.1021/ed5000718

Stöhr, C., Demazière, C., & Adawi, T. (2020). The polarizing effect of the online flipped classroom. *Computers & Education*, *147*, 103789. https://doi.org/10.1016/j.compedu.2019.103789

Tackett, J. L., & Willson, V. L. (2018). A propensity score matching evaluation of the effectiveness of a supplemental instruction program for introductory accounting courses. Journal of Education for Business, 93(4), 143-151. doi: 10.1080/08832323.2017.1423181

Talanquer, V. (2007). Explanations and Teleology in Chemistry Education. *International Journal of Science Education*, *29*(7), 853–870. https://doi.org/10.1080/09500690601087632

Talanquer, V. (2013). Chemistry Education: Ten Facets To Shape Us. *Journal of Chemical Education*, *90*(7), 832–838. https://doi.org/10.1021/ed300881v

Tatli, Z., & Ayas, A. (2010). Virtual laboratory applications in chemistry education. *Procedia - Social and Behavioral Sciences*, *9*, 938–942. https://doi.org/10.1016/j.sbspro.2010.12.263

Tawfik, A. A., Reeves, T. D., Stich, A. E., Gill, A., Hong, C., McDade, J., Pillutla, V. S., Zhou, X., & Giabbanelli, P. J. (2017). The nature and level of learner–learner interaction in a chemistry massive open online course (MOOC). *Journal of Computing in Higher Education*, *29*(3), 411–431. https://doi.org/10.1007/s12528-017-9135-3

Tough, A. M. (1971). *The adults' learning projects: A fresh approach to theory and practice in adult education*. Toronto, ON: The Ontario Institute for Studies in Education. Retrieved 30 May 2019 from http://ieti.org/tough/books/alp.htm.

Towns, M. H. (2001). Kolb for chemists: David A. Kolb and experiential learning theory.

van Gigch, J. P. (2002). Comparing the epistemologies of scientific disciplines in two distinct domains: Modern physics versus social sciences. I: The epistemology and knowledge characteristics of the physical sciences. *Systems Research and Behavioral Science*, *19*(3), 199–209. https://doi.org/10.1002/sres.465

Velardi, S., Leahy, J., Collum, K., McGuire, J., & Ladenheim, M. (2021). Adult learning theory principles in knowledge exchange networks among maple syrup producers and beekeepers in Maine. *The Journal of Agricultural Education and Extension*, *27*(1), 3-20.

Vonderwell, S. (2003). An examination of asynchronous communication experiences and perspectives of students in an online course: A case study. *The Internet and Higher Education*, *6*(1), 77–90. https://doi.org/10.1016/S1096-7516(02)00164-1

Wahyudi, G. S., Artini, L. P., & Padmadewi, N. N. (2021). SELF-DIRECTED LEARNING IN EFL DURING COVID-19 PANDEMIC: AN ANALYSIS OF TEACHER'S PERCEPTIONS AND STUDENTS'LEARNING AUTONOMY. *International Journal of Language and Literature*, *5*(2), 93-104.

Wang, R., Lowe, R., Newton, S., & Kocaturk, T. (2020). Task complexity and learning styles in situated virtual learning environments for construction higher education. *Automation in Construction*, *113*, 103148.

Wavle, S., & Ozogul, G. (2019). Investigating the Impact of Online Classes on Undergraduate Degree Completion. *Online Learning*, *23*(4). https://doi.org/10.24059/olj.v23i4.1558

Woldeamanuel, M. M., Atagana, H., & Engida, T. (2014). What makes chemistry difficult?. *African Journal of Chemical Education*, *4*(2), 31-43.

Wozniak, K. (2020). Personalized learning for adults: An emerging andragogy. *Emerging technologies and pedagogies in the curriculum*, 185-198.

Williams, A. J., & Pence, H. E. (2011). Smart Phones, a Powerful Tool in the Chemistry Classroom. *Journal of Chemical Education*, *88*(6), 683–686. https://doi.org/10.1021/ed200029p

Wu, H.-K. (2014). *Exploring visuospatial thinking in chemistry learning*. https://doi.org/10.1002/sce.10126

Yeoman, P., & Wilson, S. (2019). Designing for situated learning: Understanding the relations between material properties, designed form and emergent learning activity. *British Journal of Educational Technology*, *50*(5), 2090-2108.

Zoller, U. (2001). ALTERNATIVE ASSESSMENT AS (CRITICAL) MEANS OF FACILITATING HOCS-PROMOTING TEACHING AND LEARNING IN CHEMISTRY EDUCATION. *Chem. Educ. Res. Pract.*, *2*(1), 9–17. https://doi.org/10.1039/B1RP90004H