An Exploration of How Congruence Between Students and Their Academic Environment Impacts Student Success

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Christian Spencer Lockhart

May 2025

DEDICATION

I dedicate this dissertation to my parents, Bobby and Nikki Lockhart. From the very beginning, they surrounded me with higher education, instilling in me a deep and unwavering confidence that I belonged in academia. The heights I reach are only made possible by their love and support.

ACKNOWLEDGEMENTS

The dissertation writing process was humbling, and I am deeply grateful for the grace and patience Dr. Ishitani provided throughout this journey. Thank you for your knowledge, kindness, and humor. I would also like to thank the other members of my committee: Drs. Biddix, Gibbons, and Boyd. Each of you brought valuable insights and unique perspectives that significantly enriched my work and contributed to the final product.

My colleagues at Pellissippi State also deserve acknowledgement as many of them showed vested interest in my journey. A special thanks to Dr. Rachael Cragle who welcomed my many visits to her office to discuss both my successes and setbacks. Finally, I would like to thank my mentor and professional role model, Dr. Elizabeth Firestone. Although you are no longer my supervisor, the guidance you gave me both professionally and academically has been instrumental in helping me become who I am today. If I ever have wise moments, they are usually prefaced by, "As my former supervisor, Dr. Firestone, would say…" I will always cherish the opportunities and mentorship you provided me.

ABSTRACT

This study examined the relationship between person-environment fit (P-E fit) and student outcomes at public, two-year colleges (PTYCs) using Holland's (1997) theory of vocational personalities and work environments as the framework for defining "congruence." The research was conducted at Pellissippi State Community College, a large PTYC in the southeastern United States, with a sample of 1,257 students who completed the Interest Profiler Short Form (IP Short Form) during enrollment. Two analyses were conducted to evaluate the role of congruence in student success: (1) a sequential multiple linear regression to test its predictive value for first-semester GPA, and (2) a sequential logistic regression to assess its impact on fall-to-fall retention. Cultural and precollege characteristics were included in the analyses to examine whether congruence predicted these outcomes beyond these variables. Congruence scores were calculated by assigning RIASEC types to both students and their courses and applying the C Index to generate fit scores, which were then averaged across classes. The findings revealed that congruence was not a significant predictor of either first-semester GPA or fall-to-fall retention, suggesting that other factors may play a more critical role in influencing these outcomes. This study contributes to the limited research on PTYC student success and offers insights into the applicability of Holland's theory and P-E fit frameworks in higher education research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND STATEMENT OF THE PROBLEM

Introduction

Public two-year colleges (PTYCs) are a significant and essential part of the United States' higher education system. Data from the U.S. Department of Education's (2018, 2019) National Postsecondary Student Aid Study (NPSAS) and the Beginning Postsecondary Students (BPS) study demonstrate the scope of PTYC's role in educating U.S. students after high school. For the 2015-16 academic year, 39% of all undergraduate students attended a PTYC (NPSAS:16, 2018). Furthermore, these students are "fundamentally different" than those who enroll at four-year institutions as evidenced by the fact that PTYCs serve a larger percentage of underrepresented and vulnerable populations, including first-generation, minority, post-traditional, and part-time students when compared to other institutional types (Davidson & Wilson, 2017, p. 517). For example, in the same 2015-16 academic year, 46% of all Hispanic or Latino students, 46% of all post-traditional students, 47% percent of all first-generation students, 41% of all students in the bottom income quartile, and 63.2% of all exclusively part-time students attended PTYCs (NPSAS:16, 2018). These students also enter into higher education less academically prepared and with greater diversity in goals. These goals range from acquiring a few new skills that will make them more employable to preparing for graduate school (Davidson & Wilson, 2017). All of this information points to just how significant PTYCs are to the health of the country's higher education system.

This significance is even greater in Tennessee where PTYCs have been positioned to play an essential role in preparing the workforce for the state's future economic needs. In 2013, former TN Governor Bill Haslam launched the "Drive to 55" initiative. The titular "55" was a reference to the prediction that at least 55% of the state's population would need a postsecondary credential of some kind to keep up with job demand. However, only 32% of TN citizens had a certificate or degree beyond high school at that time. To bridge this gap, Haslam committed to establishing programs to increase educational attainment and emphasized the important role that PTYCs would play ("Governor Bill Haslam," 2013). The two most significant programs to emerge from this commitment were TN Promise and TN Reconnect. TN Promise launched in 2015 and covers college tuition for high school seniors who graduate from an eligible TN high school and enroll in a PTYC immediately after graduating high school (tnAchieves, n.d.). TN Reconnect went into effect in 2018 and, similar to TN Promise, offers free tuition to "adult" students, defined as 24 or older, who do not have a postsecondary degree and enroll in a PTYC (Tennessee Reconnect, n.d.).

Investment in Tennessee PTYC students did not end with these two flagship programs. In 2019, the Haslam Family Foundation announced a $6.2 million investment in "Knox Promise." Knox Promise is being described as a pilot program that "aims to increase the number of Knox County college graduates by providing need-based grants, a textbook stipend, a coach to support students in their college career and a summer support program" to TN Promise-eligible students (Kast, 2019). This commitment has not been ignored by politicians in other states. Since the inception of TN Promise, fifteen additional states have launched "free college" programs. While these "Promise programs" vary in design, funding, and eligibility requirements, they nonetheless represent a trend across the country that does not appear to be dissipating with the total investment going up an average of $107 million per year since 2014 (Mishory & Granville, 2019). With PTYCs playing such an increasingly significant role in the postsecondary education of the country's population, the success of these students is essential. However, outcomes for PTYC students appear to be troubling.

PTYCs graduated just 25% of the first-time, full-time credential-seeking students within three years of their initial enrollment for the 2014 entry year cohort. Some acknowledgment should be given to the fact that many of the students in this cohort transferred without completing a credential. Of the approximately 58% of students who had neither graduated nor transferred after three years, only 14% were still enrolled at these institutions. Thus, 44% of the total cohort were no longer taking classes at all (NCES, 2019). A closer look reveals that most of these dropouts occur during or after the first year.

First-year retention rates at PTYCs are low when compared to their 4-year counterparts. For first-time, full-time students, PTYCs had a first-year retention rate of 62% for the 2015-2016 academic year, whereas public 4-year institutions had a retention rate of 81% for the same period (McFarland et al., 2018). This discrepancy becomes even more apparent when looking at subgroups based on culture and other student characteristics. For example, the overall first-year retention rate drops to 49% when part-time students are included (NSC Research Center, 2018). This is particularly troubling when you consider that PTYCs enroll more part-time students than four-year institutions (Davidson & Wilson, 2017). Studies have also shown that both black and first-generation students drop out at higher rates when compared to other races/ethnicities and continuing-education students respectively (Cataldi, Bennet, & Chen, 2018; NSC Research Center, 2018).

It can be argued, understandably so, that it is unfair to compare these PTYCs to four-year institutions and that differences in admissions policies, pre-enrollment characteristics, and diversity in student goals play a large role in explaining the discrepancies in student success outcomes (Cohen et al., 2013). Nonetheless, it is important to explore the additional reasons that might explain why almost half of all students who start at public PTYCs leave higher education altogether within three years. Unfortunately, the research on student persistence at these institutions has "lacked the depth" of those done at their four-year counterparts (Davidson & Wilson, 2017, p. 527). The profile of the average PTYC student does not match that of the average four-year institution student. As a result, the findings of studies conducted in four-year settings are not easily generalizable to PTYCs due to the significant differences in both how these institutions are designed and the students they serve. Thus, there is a need for research on student departure at PTYCs (Braxton etal., 2004).

Braxton, Hirschy, and McClendon (2004) described the issue of college student departure as being an "ill-structured problem" in that it defies "a single solution" and requires a "multidisciplinary approach" (p. 1-2). The literature on the topic reflects this in the diversity of theoretical frameworks that have been used to try and understand not only student departure but also the broader topic of how the college experience changes the student. These include psychosocial, cognitive-structural, sociological, typological, and person-environment interaction theories (Pascarella & Terenzini, 2005, p. 17-61). While insights and evidenced-based practices have certainly emerged from this eclectic range of theories and models, the "ill-structured problem" remains far from solved, and there is an ongoing to need to apply novel theoretical frameworks to better understand the factors that influence student departure.

Person-environment fit (P-E fit) theories have emerged as a strong choice for guiding future research on student departure. Feldman, Smart, and Ethington (2008) argue that contemporary research on student success has largely focused on students' characteristics and behavior. Using a phrase coined by Pascarella and Terenzini (1991, 2005), Feldman et al. (2008)

described this trend as the "psychological research paradigm" (p. 367). While they do acknowledge the value in these research agendas and the support that comes from them, they believe that this has come at the cost of a decrease in research on the influence of academic environments on student success. They assert the belief that a "basic understanding of student success requires attention to both the predispositions and behaviors of college students *and* the nature of campus environments" (p. 370).

If this is true for all of higher education, then it is especially true for research on community colleges. In searching the databases for articles related to both "community colleges" and "person-environment fit," the most relevant result is Schuetz (2005). This article is not a research study. Rather, it is a call to action aimed at community college researchers to explore the influence of environmental variables on student attrition. Echoing Feldman et al, (2008), Schuetz states that defining attrition as "simply a student-driven phenomenon rather than an interaction between the student and the college environment is demonstrably false" (p. 63). In their own review of the literature on community college student dropout, Davidson and Wilson (2017) critique the "deficit model" that has permeated through research on student departure at community colleges. Here, "the *problem* with community college persistence rates has fallen on the student" (p. 518, emphasis in original). The authors echo Feldman et al. (2008) and Schuetz (2005) and assert that this has come at the cost of ignoring the campus environment.

Taken together, these articles demonstrate that higher education research has largely focused on the "person" rather than the "environment" and that there is a need for studies that seek to balance these equally important aspects of fit, especially when it comes to community college students. This study is a response to that need by applying one of the most widely researched P-E fit theoretical frameworks to explore student departure at PTYCs, Holland's (1997) theory of vocational personalities and work environments.

Statement of the Problem

PTYCs serve a significant number of postsecondary students in the United States. However, the graduation rates for these institutions are concerning. Data shows that the majority of student departure occurs during or after the first academic year. A review of the literature shows that not enough research is being conducted on this problem at PTYCs. Furthermore, there is a need to apply novel theoretical frameworks as the problem itself is complicated and must be understood from multiple perspectives. Thus, this study seeks to investigate the contributing factors to first-year student departure at PTYCs through the lens of an underutilized, yet well-established theoretical framework.

Purpose of the Study

The purpose of this study is to better understand the contributing factors to first-year student performance and persistence at a PTYC. This will be accomplished by exploring if John Holland's (1997) concept of congruence is a significant predictor of these outcomes.

Research Questions

Two research questions are addressed in this study:

1. Does congruence between students' occupational personality type and their first-semester academic environment predict first-semester GPA?
2. Does congruence between students' occupational personality type and their first-semester academic environment predict persistence into the second year?

Theoretical Framework

Holland's (1997) theory of vocational personalities and work environments will be the theoretical framework that guides this study. This theory falls within the broader category of person-environment fit (P-E fit) theories. P-E fit theories share the common underlying assumption that human behavior is the result of the interaction between both the individual and their environment and that an individual's characteristics can be matched to those of the environment, creating "fit." Fit, also referred to as "congruence," is believed to lead to a variety of desirable outcomes that have historically been the focus of much of higher education research. These outcomes include increased satisfaction, performance, and stability ("Person-Environment Fit," 2004). Holland's conceptualization of how congruence can be determined requires an understanding of its central concepts.

Holland's (1997) theory of vocational personalities and work environments has been described as the most "widely researched theory on vocational interests" (Nye, Su, Rounds, & Drasgow, 2012, p. 385). It asserts that both individuals and environments can be described in terms of their resemblance to each of six model types: (1) Realistic, (2) Investigative, (3) Artistic, (4) Social, (5) Enterprising, and (6) Conventional. Hereafter, these six types, when discussed as a whole, will be referred to as the RIASEC types. For an individual's personality, each type represents a cluster of interrelated attitudes, interests, and competencies. For the environment, each type represents the activities, tasks, and roles that are required in that setting. The relationships between these types are best represented using a hexagonal model with each type placed at a corner in the order indicated by "RIASEC" (see **Figure 1.1**).

Figure 1.1
Holland Hexagon
Note. Adapted from *Purdue University Career Success* **(**https://www.cco.purdue.edu/students/hollandhexagon**).**

In Holland's (1997) theory, congruence is determined by comparing an individual's RIASEC type to that of the environment. For example, Artistic types tend to prefer activities that do not have a specific procedure and allow for expression of creativity. Thus, they would theoretically have low congruence with a Conventional environment that requires set procedures and working with data. In a postsecondary setting, "the persistence, satisfaction, and achievement of students are functions of the congruence or fit between students and their academic environments" (Smart et al., 2000, p. 2). Several prominent and well researched assessments have been developed to provide the individual with an overview of their personality based on Holland's theory. These assessments typically provide the individual with their personality profile, ordered from their highest RIASEC type to lowest, based on their answers to items that correspond to each type. A thorough overview of these assessments, including the one that will be used in this study, will be provided in Chapter Two.

RIASEC types are assigned to an environment in one of two ways. The first is by assessing the RIASEC profiles of the individuals who occupy that environment, and the second is by analyzing the skills, knowledge, and tasks required in the environment (Gottfredson & Holland, 1996). Over the last 50 years, researchers have developed a variety of methods for calculating congruence scores using the RIASEC profile for both the individual and the environment. A review of these methods, including the one that will be used in this study, will also be provided in Chapter Two.

Despite the volume of research that has been conducted in other fields using Holland's theory, it is rarely applied in higher education studies. Smart, Feldman, and Ethington (2000) stated that there is a "virtual lack of attention to Holland's theory by higher education scholars" (p. 32). This deficit can likely be attributed to two final points about Holland's (1997) theory, its relevance to higher education, and the potential contributions of this study. The first is that contemporary research has largely focused on students' characteristics and behavior and that not enough attention has been paid to the academic environment (Feldman, Smart, & Ethington, 2008). Thus, there is a need for a theoretical approach that places renewed emphasis on the role of the environment. Second, it has been observed that there are unresolved methodological issues when attempting to define what constitutes the "academic environment" that has limited the applicability of Holland's theory and the generalizability of results when it is used (Allen & Robbins, 2008; Tracy & Robbins, 2006). This study will define the academic environment in a novel way in order to better evaluate the theory's effectiveness in higher education settings. Both of these points will be discussed further in Chapter Two.

Significance of the Study

This study is significant for several reasons. First, it will attempt to contribute to the field's understanding of the "ill-structured problem" that is college student departure. Second, it seeks to understand this problem in a setting that is both incredibly important to the health of higher education in the United States and under-researched, the public two-year college. Third, this study will do its part to answer calls from the literature to look at student departure as the result of an interaction between both the student *and* the academic environment. This will be accomplished by applying one of the most widely researched theoretical frameworks for understanding the interaction between an individual and their environment, Holland's (1997) theory of vocational personalities and work environments. While a small number of researchers have already attempted to apply this theoretical framework, there has been a methodological issue when attempting to define how the "academic environment" should be defined. Thus, the final contribution of this study will be to use a novel methodological approach that reconceptualizes the academic environment and how it is measured in a way that could increase the applicability of Holland's theory to college students.

Terminology

Personality and Environmental Patterns: Holland (1997) defines a person's personality pattern as their "profile of resemblances to the personality types" (p 31). Personality patterns can include as few as two of the personality types and as many as all six and are traditionally represented using the first letter of each type, ordered from most to least resemblance. For example, if an individual has a personality pattern of SACREI, the implication is that their preferred activities, interests, abilities, and values most closely resemble those associated with the Social personality type and least closely resemble those of the Investigative personality type. Likewise, environmental patterns describe how closely a particular environment resembles the six model environments and are also represented by rank-ordering each type from highest to lowest using the first letter of each type. Researchers and practitioners have developed a variety of methods and instruments for determining both the personality and environmental patterns of people and environments, respectively. These will be detailed in Chapter Two.

Holland Code: When a personality or environmental pattern includes only the top three types, once again represented by using only the first letters, this is traditionally referred to as a Holland Code. Using the earlier example, a person with a full personality pattern of SACREI would have SAC as their Holland Code. In both research and application, the Holland Code is by far the most common way of representing both personality and environmental patterns. A person's first type is typically called their dominant type, and the second and third types are referred to as the secondary and tertiary types (Nauta, 2013).

Congruence: Nuata (2013) defines Holland's conceptualization of congruence as "the degree of fit between an individual and his or her current or projected environment with respect to the RIASEC types" (p. 60). In other words, the more similar an individual's personality pattern is to the pattern of an environment, the greater the congruence. Chapter Two will include an overview of prominent methods for calculating a congruence score.

Traditional and Post-Traditional Students: While age is a continuous concept, it is a common practice in higher education research to categorize students based on their age at enrollment in order to compare the outcomes of these students. Students who fall within the age range of an individual who enrolled in a postsecondary institution immediately after high school and then progressed to complete their undergraduate degree in four years are typically operationalized as “traditional” students (Goodman, 2024; Hoachlander, Sikora, & Horn, 2003; Spencer, de Novais, Chen-Bendle, & Ndika, 2024).

While there have been a variety of terms to describe students who fall outside of this enrollment trajectory, recent literature has used “post-traditional” as an inclusive descriptor for these students. While it does not exclusively refer to age, post-traditional students are “typically adults in their mid-20s and over” (Spencer, de Novais, Chen-Bendle, & Ndika, 2024). For the purpose of this study, post-traditional students will be used primarily to refer to students who are 24 years of age or older. This cutoff point was chosen as it is consistent with the eligibility requirements for the previously discussed TN Reconnect program. Thus, traditional students will be defined as students who are 23 years of age or younger.

Organization of the Study

This study is organized into five chapters. Chapter One provided an overview of the need for the study, its purpose, research questions, theoretical framework, and significance. Relevant literature is critically reviewed in Chapter Two. Chapter Three details the methods and procedures that are used in the study, and the findings of the study are described in Chapter Four. Finally, Chapter Five provides an overall summary of the study including a summary and discussion of the findings, final conclusions, and recommendations for future research.

CHAPTER 2

LITERATURE REVIEW

Introduction

The purpose of this research was to better understand the contributing factors to first-year student departure at a large, public community college by looking at the relationship between John Holland's (1997) concept of congruence and indicators of academic success, first-semester GPA and persistence into the second academic year. The following literature review is divided into three sections. The first section discusses student persistence at public two-year colleges (PTYCs) by highlighting the applicability of Tinto's (1975, 1986, 1993) interactionalist theory of student departure to PTYCs before transitioning into a how a P-E fit framework can overcome its shortcomings. The second section focuses specifically on Holland's (1997) theory and includes an overview of methods for classifying people, environments, and calculating congruence. Finally, the third section will review relevant cultural and demographic characteristics that should be considered when applying Holland's theory and their relevance to student outcomes at PTYCs.

Student Persistence at Public Two-Year Colleges

Tinto's Interactionalist Theory of Student Departure

Research on student departure for much of the last 45 years has largely been guided by Tinto's (1975, 1986, 1993) interactionalist theory of college student departure. Pascarella and Terenzini (2005) describe his theory as "probably the most widely used framework guiding research into the complex persistence-related interconnections among students and their college experiences," and Braxton, Hirschy, and McClendon (2004) assert that it "holds paradigmatic status as a framework for understanding student departure (p. 425; p. 2). Tinto's theory is based on the work of sociologist Emile Durkheim (1951) and posits that a student persists at an institution when they achieve social and academic integration (Tinto, 1993). **Figure 2.1** shows Tinto's (1993) complete model. As the model depicts, academic and social integration are theorized to result from the interaction between a student's attributes and commitments and their institutional experiences. Different types of experiences are expected to contribute to each type of integration. Social integration is believed to ensue from extracurricular activities and peer group interactions and academic integration from academic performance and faculty/staff interactions.

While the volume is nowhere near what has been done for four-year institutions, a modest number of researchers have applied Tinto's (1975, 1986, 1993) model to the study of student persistence at PTYCs over the last forty years. Tinto (1993), himself, questioned if his

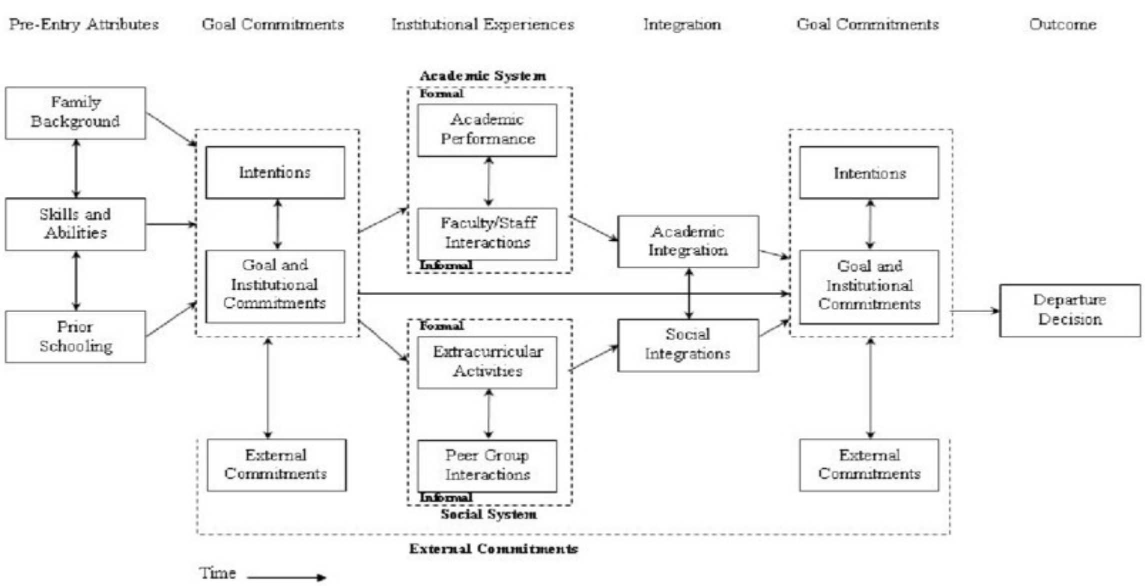

Figure 2.1
Tinto's (1993) Longitudinal Model of Institutional Departure

theory was applicable to the two-year setting and stated that these students are "likely to experience a wide range of competing external pressures on their time and energies and to be unable to spend significant amounts of time on campus interacting with other students and members of the staff" (p. 78). Some research in the field has validated Tinto's concerns about using his model for two-year institutions.

Voorhees (1987) used Tinto's (1975) model as his theoretical orientation to explore the persistence of 369 new and continuing students enrolled at a single two-year college during the fall of 1984. Persistence was defined as returning in one or both of the following two semesters. Student intention to return, purpose for enrolling, demographic characteristics, and data representing academic integration were used as independent variables in a log-linear analysis. Gender, purpose for enrolling, and intent to return were found to be significant predictors. Variables related to academic integration (i.e., GPA, interactions with faculty, and number of hours spent studying) were not found to be significant. Based on his findings, the author concluded that academic integration "may be of less importance in explaining the persistence of community college students… than four-year college or university students" (p. 127).

Kubala and Borglum (2000) came to a similar conclusion after conducting their own single site study with 462 community college students. The authors collected survey data on 53 items related to pre-entry attributes, goals and intentions, social integration, and academic integration. Items related to social integration included time spent on campus and involvement with campus activities. Academic integration included hours per week spent studying and satisfaction with the availability of instructors. Using one-way analyses of variance, the author found no significant relationship between withdrawal rates and either academic or social integration.

Other studies have come to more optimistic conclusions about the applicability of Tinto's (1975, 1986, 1993) theory to PTYC students. Pascarella and Chapman (1983) included two-year students alongside four-students in their investigation of the validity of Tinto's (1975) model using path analysis. The sample was comprised of 2,326 students and included longitudinal data across two consecutive academic years. Student background and institutional characteristics were included in the model alongside students' answers across two scales designed to operationalize academic and social integration. They found that, unlike students at residential universities where persistence was more strongly influenced by social integration, students at 2-year commuter institutions were "defined by a significant degree by academic integration" (p. 98).

Pascarella, Smart, and Ethington (1986) used Tinto's concepts to look at the long-term persistence and withdrawal behaviors of students at two-year institutions using a sample of 825 students from 85 institutions. Data on these students came from two surveys that include five independent variables based on Tinto's (1975) theory: (1) background characteristics, (2) initial commitments, (3) academic and social integration, (4) subsequent goal and institutional commitments, and (5) persistence/withdrawal behaviors. Degree persistence and degree completion were the dependent variables. Structural equations were used to determine if Tinto's (1975) theory predicted degree persistence and completion. The author concluded that "the results tend to confirm the importance of person-environment fit as a salient influence on degree persistence and completion in postsecondary education" (p. 47).

These limited, mixed, and dated results are far from conclusive. In their own review of Tinto's (1975, 1986, 1993) theory, Braxton, Hirschy, and McClendon (2004) concluded that "the explanatory power of Tinto's theory to account for student departure in two-year college remains undetermined and open to empirical treatment" (p. 18). In the absence of conclusive quantitative findings, qualitative studies can often provide additional insight into a topic. There have been some qualitative studies that have sought to better understand the PTYC student experience through the lens of academic and social integration.

Karp, Hughes, and O'Gara (2010) applied Tinto's (1993) theory to community college students using an exploratory, qualitative method. They conducted two semi-structured interviews six months apart with 36 students. The first interview explored students' initial experiences in college, and the second interview focused on the decision to continue or not with their enrollment. Interviews were read thematically, using Tinto's (1993) theory as a theoretical lens. Specifically, they looked for themes related to integration, operationalized as "sense of belonging" (p. 71). They found that 70% of the students reported having a sense of belonging and that 90% of these had persisted into the second year. A salient theme that emerged across these students was the importance of social relationships that served to enhance their institutional knowledge. The authors labeled these relationships "information networks" (p. 76).

Deil-Amen (2011) also employed a qualitative method to explore if Tinto's theory was applicable for understanding the persistence of students at two-year institutions by conducting 238 semi-structured interviews with students, staff, and faculty across seven public and seven private two-year institutions. Of the 125 students who were interviewed, 92% indicated that it was a staff or faculty member at the institution who had been "instrumental to their sense of adjustment, comfort, belonging, and competence as college students" (p. 61). Like Karp et al. (2010), Deil-Amen found that in-class experiences play a big part in integration and that social relationships are often based in "academic utility" causing them to coin the term "socio-academic integrative moments" to described student experiences where the line between academic and social integration becomes blurred (p. 82). When asked to describe what led to their persistence, it was these socio-academic moments that students most often articulated.

Reviewing this literature holistically, two themes emerge. First and most obvious, there is a desperate need for more research on PTYC student persistence. Second, the most influential model for understanding student persistence, Tinto's (1993) interactionalist theory of student departure, does not appear to be a totally effective theoretical framework for researching this problem. These studies suggest that this is likely due to the often-diminished role that social integration seems to play at PTYCs. In a field that has focused its research heavily on social integration and its related terms such as connectedness, community, support, relatedness, and, most popular of all, belonging, perhaps higher education researchers have not yet developed or found a theory through which PTYC persistence can be accurately understood.

Person-Environment Fit

As was cited in Chapter One, the "ill-structured" problem of college student departure cannot be solved with a single solution. Rather, a "multidisciplinary approach is needed' (Braxton et al., 2004, p. 1-2). Stuart, Rios-Aguilar, & Deil-Amen (2014) echo this sentiment when they assert that "there is a need for an alternative model of student persistence to fill in some of the gaps not addressed by Tinto's framework" (p. 330). Inspiration for where one might look to find something to fill these "gaps" can found in the urgings of prominent authors who have argued that much of higher education research has focused disproportionately on student characteristics and that equal attention must be given to the "nature of campus environments" (Feldman et al., 2008, p. 370). Schuetz (2005) and Davidson and Wilson (2017) are equally critical that the majority of community college research that has disregarded the role of the academic environment in favor of placing the preponderance of responsibility on the student for persisting or departing.

Thus, person-environment fit theories (P-E fit) emerge as area ripe for exploration. If understanding college student outcomes requires a multidisciplinary solution, then the "multidimensional" nature of P-E fit may meet the call (Etzel & Nagy, 2021). Numerous authors have observed that Tinto's model is, itself, a theory of person-environment fit (Pascarella et al, 1986; Schuetz, 2005; Smart, 2000). However, a review of his model in **Figure 2.1** with these previously stated criticisms in mind points to a potential issue in its design. There is no place in it for characteristics of the academic environment and the potential to highlight the congruence or incongruence between those characteristics and the attributes of the student. There is only the latter and the institutional experiences that may or may not occur. In other words, there is a gap, and P-E Fit theories offer a potential bridge. This bridge is flexible and can be used in a variety of research contexts as long as there is an attribute of the student and a corresponding characteristic of the environment with which you wish to determine fit or congruence. **Figure 2.2** provides a simple example of what Tinto's model might look like if it were revised to consider P-E fit more explicitly.

As an umbrella term, P-E fit encompasses a range of theories, but these can be broken down into a few subcategories. First, Edwards, Cable, Williamson, Lamber, and Shipp (2006) distinguish between three types of P-E fit: 1) needs-supplies fit (N-S fit), 2) demands-abilities fit (D-A fit), and 3) supplementary fit. Needs-supplies fit is the "comparison between the psychological needs (i.e., desires, values, goals) of the person and the environmental supplies

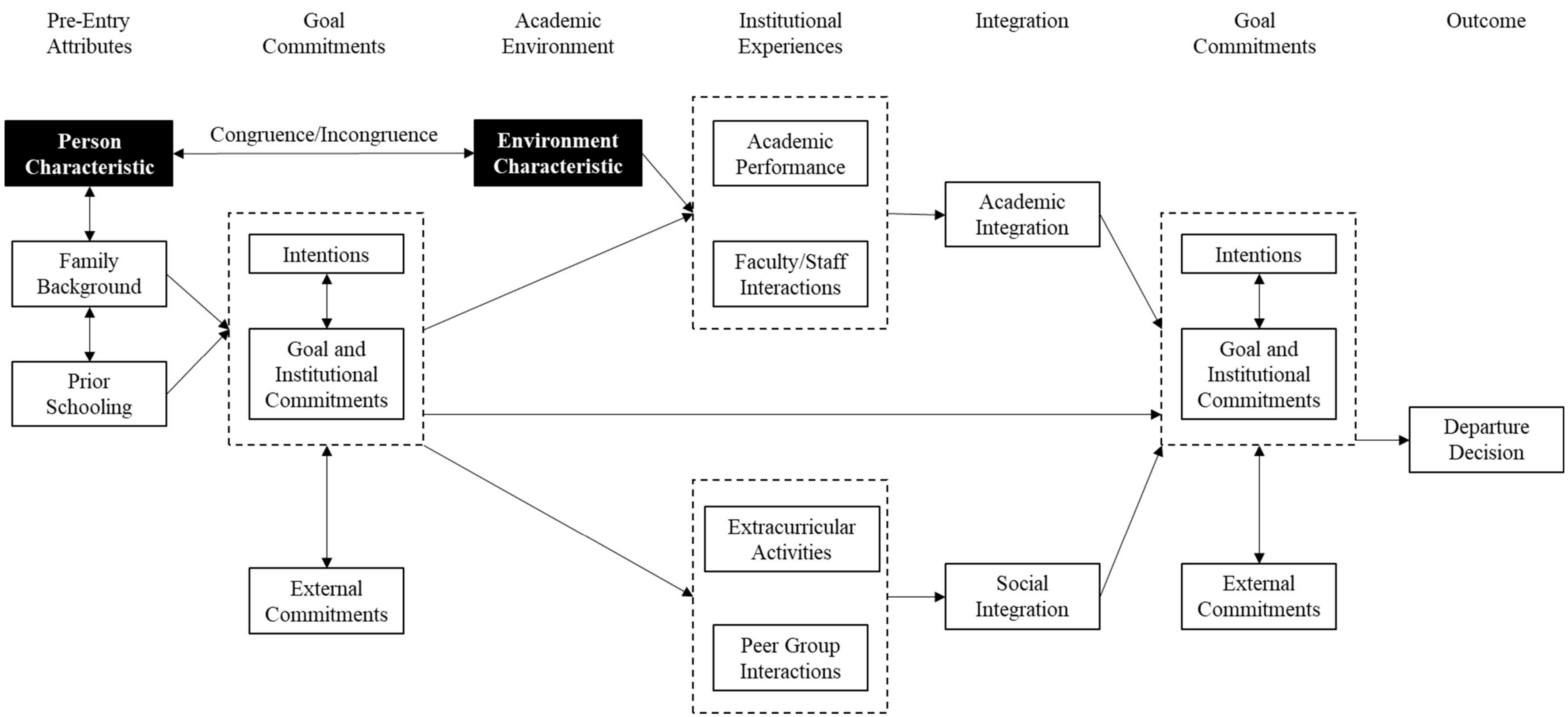

Figure 2.2
Tinto's (1993) Model Revised to Include P-E Fit

that serve as rewards for needs" (p. 804). Demands-abilities fit refers to the match between the individual's knowledge, skills, and aptitudes with demands of the environment. Supplementary fit is the similarity between the person and the individuals within that environment.

Edwards and colleagues go on to elaborate that there are also three kinds of P-E fit studies: 1) atomistic, 2) molecular, and 3) molar. Atomistic studies are those that measure characteristics of the person and the environment separately and then combine them to represent fit. The molecular approach entails that researchers directly assess discrepancies between the person and environment. Finally, molar methods focus on the person's *perceived* fit with their environment. Given the range of concepts and methodologies that can fall within the types of fit and fit research respectively, it is of no surprise that there are some non-Tinto examples of higher education researchers applying P-E fit theories to understanding student experience.

Gilbreath, Kim, and Nichols (2011) looked at the relationship between student-university fit and satisfaction and well-being with a sample of 252 students across two four-year institutions. This study of needs-supplies defined the environment as the entire university and implemented. The authors generated 18 "fit items" that represent students' needs and then developed a questionnaire to measure them. Principal Axis Factor Analysis was used on the resulting data that resulted in a three-factor solution. Based on the items included in each factor, they were labeled "social environment," "academic environment," and "physical environment." For each participant, student needs and university supplies scores were calculated for each factor. These scores were then used in polynomial regression analyses that calculated change in the dependent variables based on the extent to which a university's supplies matched, exceeded, or fell below students' needs. The results supported that the fit between the students and these environmental factors predicted higher levels of satisfaction and wellbeing.

Li, Yao, Chen, and Wang (2013) explored the relationship between three different types of fit in an academic environment with a sample of 246 freshman at a four-year university. In addition to D-A fit and N-S fit, the author identified interest-major fit (I-M fit) as a separate category of fit in what they called "The Model of Three-Factor Fit Perceptions." The authors developed a 9-item Perceived Academic Fit Scale with each type of fit having three corresponding questions. They administered this along with multiple other scales and items to measure their dependent variables of interest, academic satisfaction, depression, academic performance, and major change intention. An incremental validity analysis was used to determine what each type of academic fit contributed to the overall prediction above and beyond the other two. The authors conducted a series of sequential multiple linear regressions for each of the dependent variables. Significant findings included that I-M fit predicted greater proportions of variance for major change intention, D-A fit predicted greater proportions of variance in academic performance, and N-S fit predicted greater proportions of variance in academic satisfaction and depression.

Bowman and Denson (2014) created and administered the Student-Institution Fit Instrument (SIFI) to 409 students from a large, public university in the Midwest and 99 students from a small, religiously affiliated liberal arts school completed the SIFI. The SIFI uses student responses to 7-point Likert scale questions to generate scores across eight domains of student-institution fit: (1) academic, (2) social, (3) cultural, (4) physical, (5) athletic, (6) religious, (7) socioeconomic, and (8) political. Structural equation modeling (SEM) was used to measure the relationships between the student-institution fit as a predictor variable and college satisfaction, social isolation, academic adjustment, and intent to persist as outcome variables. The results of the SEM indicated the student-institution fit is significantly and positively related to college satisfaction. Academic adjustment was only correlated with the academic dimension of the SIFI, and student-institution fit was negatively correlated with social isolation and positively related to intent to persist.

Bohndick, Rosman, Kohlmeyer, and Buhl (2018) sought out to explore the relationship between subjective D-A fit and academic success with a sample of 693 teacher education students. Each member of the sample completed a questionnaire within which they were asked to rate both their own ability and their program across several general requirements. Subjective fit was then determined by subtracting the ability ratings from the program ratings. Academic success was determined based on self-reported grades, perceived performance, and study satisfaction. The authors used structural equation modeling as their statistical analysis. Their results indicated that academic success is best achieved when abilities *match* demands rather than exceed.

Etzel and Nagy (2021) conducted a study on the relationship between a multidimensional conception of P-E fit with educational outcomes with 746 university students in Germany. They developed a 12-item measure that covers four dimensions of P-E fit: 1) interest-contents fit (I-C fit), 2) N-S fit, 3) D-A fit, and 4) values-culture fit (V-C fit). Using structural equation modeling, the authors assessed how both the overall measure and each dimension of the measure were associated with self-reported satisfaction with contents, satisfaction with conditions, drop-out intention, and academic performance. The overall measure was significantly related to both outcomes. As for each dimension, I-C fit was significant for satisfaction with study contents, N-S fit was significant for satisfaction with study conditions, and D-A fit was significant for self-reported performance. The author concluded in their discussion that they “found evidence for the importance of considering different content dimensions of fit” (p. 374).

While these studies demonstrate that looking at student outcomes through the lens of P-E fit theories is promising, they are limited to 4-year institutional settings and utilize homegrown methods of measuring student characteristics that require additional validity and reliability research. Thus, they do nothing to answer the criticisms presented by Schuetz (2008) and Davidson and Wilson (2017) that not enough research has been done on how the academic environment impacts student outcomes. It is then worth exploring the landscape of established and well researched P-E fit theories to find one that allows researchers to categorize the academic environment and determine the extent to which an assessed student characteristic is congruent with that academic environment. To that end, this literature review now turns to Holland's (1997) theory of vocational personalities and work environments

Holland's Theory of Vocational Personalities and Work Environment

General Concepts

Per Feldman, Smart, and Ethington (2004, p. 334- 335, emphasis in original), Holland's theory has three basic assumptions. The first is the "*self-selection assumption"* and asserts that individuals will naturally choose environments that align with their personality types. The second is the "*socialization assumption*" and posits that environments "require, reinforce, and reward… attitudes, interests, and competencies" that align with those that are represented in that environment. The third and most relevant to this study is the "*congruence assumption*." Here, environmental "stability, satisfaction, and achievement" are the result of fit between an individual and their environment. These authors conclude their description of these assumptions with the statement that "[v]oluminous empirical evidence generally supports the validity of each of these three fundamental assumptions."

In addition to congruence, Holland (1997) also proposed two other concepts that are relevant when attempting to apply his theory to predict the outcomes of individuals in environments: consistency and differentiation. Consistency refers to “the degree of relatedness between personality types or between environmental models” (p. 4). In other words, it is the within-profile relationship between types. Like congruence, Holland stated that there are levels of consistency between types based on their proximity on the hexagonal model. Adjacent types are the most consistent, opposite types are the least consistent, and alternate types have intermediate level of consistency. Differentiation is defined as the “magnitude of the difference between highest and lowest scores on the six variables” (p. 57). A well differentiated profile will have high peaks and low valleys whereas an undifferentiated profile will be relatively flat. Along with high congruence, Holland asserted that higher levels of consistency and differentiation lead to more predictable behavior.

Classifying People Using the Holland Types

A number of measures and strategies have been developed since the inception of Holland’s theory to classify individuals based on the RIASEC types. What follows is an overview of some of the most popular examples used both in research and practical application of the theory. A more in-depth review of the O*Net Interest Profiler, in particular, will be provided as it is the instrument used in this study.

Qualitative Methods. Feldman, Smart, and Ethington (2008) used the phrase “qualitative methods” to describe classifications that are based on self-selection assumption. If it is assumed that individuals self-select environments that are congruent with their personalities, then it should follow that an individual can be categorized by the environment that they have chosen. For example, Ishitani (2010) determined the RIASEC types of college graduates based on the major they graduated with a degree in. Similarly, Miller, Wells Springer, and Cowger (2003) assigned RIASEC types to parents based on their current occupations. These methods require an existing and established classification system for the environments and are best used in situations where the individual has demonstrated the expected outcomes of congruence (i.e., stability, satisfaction, and achievement).

Strong Interest Inventory. The Strong Interest Inventory (SII; Harmon, Hansen, Borgen, & Hammer, 1994) contains 317 items. The majority of the items contain either an activity, academic subject, occupation, or type of person and respondents choose between "like," "dislike," or "neutral." Furthermore, respondents are also asked to indicate their preference between multiple pairs of self-descriptive statements by marking "Yes" or "No." While earlier versions were paper-and-pencil, the latest iteration must be completed and scored electronically (Strong Interest Inventory, 2004).

Self-Directed Search (Form R). The Self-Directed Search (SDS; Holland 1994) is an assessment of self-reported interests and skills. Subscale scores are calculated for each of Holland's six occupational personality types (RIASEC). This instrument has demonstrated both reliability and validity in its application. Holland, Fritzche, and Powell (1994) confirmed the instruments test-retest reliability with test-retest correlations ranging from r = .76 (Conventional) to r = .89 (Artistic & Social). Holland (1997) reported that the SDS has concurrent validity.

Unisex Edition of the ACT Interest Inventory. The Unisex edition of the ACT Interest Inventory (UNIACT; ACT, 1995) is a 90-item assessment that is available to students when they register to take the ACT. Each item corresponds to a Holland type and asks participants to choose "dislike," "indifferent," or "like" in response to a work-related activity. Results are then calculated to generate a score, ranging for 20-80, for each of the six types. The score represents the level of interest for that type with higher scores indicating greater interest. The results of this assessment have been shown to be reliable (ACT, 1995).

*O*Net Interest Profiler*. The O*Net Interest Profiler (IP; Lewis & Rivkin, 1999) was originally developed by the U.S. Department of Labor as a 180-item, self-scored, paper-and-pencil interest inventory that provided assessment takers with scores for each of the RIASEC types. This initial version has come to be called the IP Long Form as newer, shorter versions were developed in order to increase ease of use. The IP Short Form (Rounds et al, 2010) was created in 2010 and contains 60 items. The Mini-IP (Rounds et al, 2016) was developed in 2016 for mobile devices and contains just 30 items. Each version has been computerized for electronic use with the IP Short Form being the most widely used due to inclusion in O*Net's My Next Move resource which averages one million visits per month. Furthermore, the IP Short Form is available for incorporation into the applications of public and private organizations for free and O*NET Web Services provides technical support on how to do so (Rounds et al, 2018).

Rounds, Su, Lewis, and Rivkin (2010) provided an overview of the development of the IP Short Form and its psychometric characteristics. The developmental sample for the IP Short Form included 1,061 participants from each region of the United Stated. Of these, 41% were male and 59% female. In terms of ethnicity, 25% were African American, 59% White non-Hispanic, 10% Hispanic, and 6% other racial/ethnic groups. The sample also included a wide range of ages with 10% being 18 years or less, 16% 19 to 22, 24% 23 to 30, 24% 31 to 40, 17% 41 to 50, and 9% 50 or older. Approximately 8% of the of participants reported being in a junior college or vocational school and approximately 14% reported they were in "college." For reliability, the authors checked internal consistency for each of the RIASEC scales and found that they ranged from .78 to .89. Furthermore, test-retest correlations across two test occasions ranged from .78 to .86 indicating that the scales are highly stable. The authors also found evidence of convergent and discriminant validity by examining the cross correlations between the IP Short Form and the Interest-Finder RIASEC scales (Wall & Baker, 1997). Same-named scale correlations ranged from .74 to .82 and dissimilar scale correlations ranged from .12 to .48. Furthermore, intercorrelations between the IP Short Form's RIASEC scales conformed to a circular order pattern, consistent with the hexagonal relationship between Holland's type and indicative of structural validity.

A point should be made about these assessments and any others that are based on Holland's theory. Holland (1997) conceptualized each of the RIASEC types as holistic models that include a set of attitudes, values, skills, abilities, and interests. For example, in addition to preferring tools, machines, and animals, Holland also described Realistic individuals as being typically conforming, genuine, and inflexible. However, most, if not all, Holland-based assessments are *interest* inventories exclusively and do not measure these other aspects of one's personality, values, skills, and abilities. Holland did assert that "vocational interests are an expression of personality" and that "interest scales are related to a person's values, academic achievement, liberalism, adventurousness, and other personal characteristics" (p. 8-9). Nonetheless, it should be clear that what is being measured is an individual's interests and that other qualities may only be indirectly inferred based on the model types as Holland has defined them.

Classifying Environments Using the Holland Types

Defining what the environment is and accurately classifying it using Holland's RIASEC typology is essential for research using his theory as a framework. A variety of methods for determining the Holland code for an environment have evolved and become more sophisticated since the inception of the theory. Rounds et al. (1999) organized these various methods into three categories: (1) the incumbent method, (2) the empirical method, and (3) the judgment method. However, these methods have almost exclusively been used to classify occupational environments. This is not to say that there have not been attempts also classify academic environments. However, the primary method for doing so is distinct from the three aforementioned categories and, for the purpose of this review, will be called the "association method." What follows is an overview of the evolution of these four methods and how they have historically been used to categorize environments using the RIASEC types.

The incumbent method is best described using Holland's (1997) assertion that "the character of an environment reflects the nature of its members" (p. 41 – 42). To put it another way, "people in the environment are the environment" (Rounds et al., 1999, p. 3). This approach to environmental categorization uses the RIASEC interest scores of the individuals who occupy an environment to determine its Holland Code. Thus, the incumbent method is the reverse of the qualitative method that was detailed in the section on classifying individuals. Whereas the qualitative method categorizes the individual based on the environment, the incumbent method categorizes the environment based the individuals that make it up. The Environmental Assessment Technique (EAT) is perhaps the most well-known example of the incumbent method. Astin and Holland (1961) originally developed the EAT as a means for classifying a college environment. It requires a census of the members in the environment and an assessment of their dominant RIASEC type. Once the number of members who belong to each RIASEC type has been determined, they are converted to percentages of the total population. For example, if thirty out of a hundred people are categorized as Social, then 30% of the total population is considered Social. The six RIASEC types are then arranged from highest percentage to lowest to create its final environmental pattern.

The 1st edition of the *Dictionary of Holland Occupational Codes* (DHOC) was published in 1982. Its development brought on a "methodological change" away from classifying environments based on the interests of its members (Gottfredson & Holland, 1996, p. 717). Instead, an empirical method was used where RIASEC types are assigned to environments by using occupational analysis data (Rounds et al., 1999, p. 3). Gottfredson, Holland, and Ogawa (1982) used multiple discriminant analysis with data from the *Dictionary of Occupational Titles* (DOT) to develop "classification equations." These equations were used to assign a Holland code to just over 12,000 occupations. Three editions of the DHOC have been published and each subsequent edition has used the same procedure to add new occupations and update old ones as needed based on new data from the DOT (Gottfredson & Holland, 1989; Gottfredson & Holland, 1996).

In addition to using these classification equations, the 3rd and most recent edition of the DHOC also applied a judgment method as a means of reviewing their empirically developed Holland codes for each occupation. Here, trained judges directly reviewed a job's description to determine the most appropriate RIASEC categorization (Rounds et al., 1999, p. 5). As part of this review, the authors checked Holland codes against other empirical sources, including data from applications of the Position Classification Inventory (PCI). Like the EAT, the PCI was developed as a method for giving a specific environment a Holland code. Rather than using the personality types of its members, however, the PCI classifies the environment based on its "demands and rewards" by scoring the results of an 84-item inventory (Gottfredson & Holland, 1991, p. 49). When there appeared to be a mismatch between the empirically and judgment derived Holland code, the job analysis data was manually inspected and discussed to determine if a revised Holland code was needed. This process resulted in the classification of 12,860 occupational titles (Gottfredson & Holland, 1996).

The 3rd edition of the DHOC also includes an index of instructional programs and corresponding Holland codes. However, the Holland codes for these environments were not developed using incumbent, empirical, or judgment methods. Rather, these instructional programs were categorized based on their association with occupational environments that had already been classified using the RIASEC types. This association method was done using data from the *Classification of Instructional Programs* (CIP; NCES, 1991). As the name indicates, the CIP is a classification system of programs offered by postsecondary institutions across the United States. It is updated approximately every ten years by the National Center for Education Statistics to include new instructional programs and fields of study (NCES, 2020). The 1990 revision of the CIP included a "crosswalk" with the DOT. A crosswalk matches items from one dataset to those in another based on an underlying principle. In this case, the occupational titles from the DOT were matched to the corresponding instructional program in the CIP. For example, the occupational title "Accountants & Auditors" was matched with the instructional program "Accounting."

These associations made it possible for Gottfredson and Holland (1996) to produce Holland codes for each instructional program based on the RIASEC types that the corresponding occupational titles had been assigned. To do this, they began by assigning a score of 0 to 3 for each RIASEC type for every occupation in the DOT based on the work they had already done. A score of 3 was assigned to the dominant type, a 2 to the secondary, a 1 to the tertiary, and a 0 to the remaining types. For example, the Holland code for "Accounting & Auditors" is CEI. **Figure 2.3** shows how it would be scored.

In most cases, a particular instructional program would be associated with multiple occupational titles. The Holland code for that program would be assigned by adding up the scores in each RIASEC category across every associated occupational title. The RIASEC category with the highest score would then be assigned as the dominant type to that instructional program, the second highest as the secondary, and the third highest as the tertiary. **Figure 2.4** provides an example. Based on the final scores, the Holland code for the Accounting instructional program would be CEI. This method resulted in the classification of the 1,457 instructional programs in the CIP using the Holland typology (Gottfredson & Holland, 1996).

To conclude, it is worth highlighting a few general points that Holland (1997) asserted when discussing the classification of environments. First, he observed that environments "rarely have a homogenous character," and as a result, one must pay attention to "where a person is located within a particular environment [to] determine the kind and amount of stimulation they will receive." To this end, he encouraged practitioners and researchers to assess a person's "subenvironments" or "subunits" when attempting to apply the concept of congruence. Other influences that must be taken into consideration and controlled for include "time spent" and the "distribution of power" in an environment (Holland, 1997, p. 42 - 43). These points will be revisited in the discussion of congruence research in higher education and how this study will conceptualize and categorize the academic environment.

Occupational Title	R	I	A	S	E	C
Accounting & Auditors	0	1	0	0	2	3

Figure 2.3
Profile Score for Accounting & Auditors

Instructional Program: Accounting						
Occupational Title	**R**	**I**	**A**	**S**	**E**	**C**
Accounting & Auditors	0	1	0	0	2	3
Budget Analysts	0	1	0	0	2	3
Credit Analysts	0	0	0	0	2	3
Financial Examiners	0	0	0	0	3	2
Tax Examiners, Collectors, & Revenue Agents	0	0	0	0	2	3
	0	**2**	**0**	**0**	**11**	**14**

Figure 2.4
Profile Score for Accounting Instructional Program

Calculating Congruence

Operationalizing congruence has been an ongoing issue since Holland introduced the concept and researchers began using it in their studies. In fact, multiple authors have asserted that issues in operationalizing congruence call into question past attempts at testing it in studies (Brown and Gore, 1994; Camp & Chartrand, 1992; Spokane, 1985; Spokane, Meir, & Catalano, 2000). Thus, choosing an appropriate calculation of congruence is essential to successfully applying Holland's theory. This section reviews prominent methods of calculating congruence.

Letter agreement indices. The earliest and simplest index was proposed by Holland (1963) and was based entirely on the congruence between the first letter of the individual's and the environment's respective codes. If they were identical, the individual would be categorized as "congruent" and, if not, "incongruent." This resulted in a dichotomous variable. Holland (1973) would go on to revise this index to include consideration of hexagonal distance between the first letters. Opposite letters are assigned a value of 1, alternate letters a 2, adjacent letters a 3, and identical letters a 4. Other researchers attempted to further improve calculations of congruence by including the first two letters (Healy & Mourton, 1983) or all three letters (Gati, 1985; Iachan, 1984; Kwak & Pulvino, 1982; Robbins, Thomas, Harvey, & Kandefer, 1978 Wolfe & Betz, 1981; Zener & Schnuelle, 1976).

Camp and Chartrand (1992) applied 13 different congruence indices to the same participants and measures in order to compare them with one another by calculating intercorrelations. They found significant variability with correlations between the various indices ranging from r = .05 to r. = .98. Thus, while they purport to measure the same construct, these measures are not interchangeable, and some are likely unreliable measures of congruence. Brown and Gore (1994) continued this review of congruence indices by using simulated data to compare 10 congruence measures. They found a variety of issues and concluded that "the results, when considered together, appeared to converge on a single rather unequivocal truth; namely, that congruence indices are not all created equal" (p. 324). In response to this and the need for a congruence index that is normally distributed and takes into consideration all three types and the order in which they appear, they developed the C Index which will be covered in greater detail below.

C Index. Here, an individual's top three types are compared to the corresponding top three types of the environment in order to determine the hexagonal distance between each letter. A value of 3 is assigned if the two types are a perfect match (Artistic – Artistic), 2 is assigned if the types are adjacent on the hexagon (Artistic – Investigative/Social), 1 if they are two spaces apart (Artistic – Realistic/Enterprising), and 0 if they are on opposite sides of the hexagon (Artistic – Conventional). For example, consider an individual with the Holland code SAE and an environment with the code SIR. The first value would be 3 because there is a perfect match, the second value would be 2 because Artistic and Investigative are adjacent on the hexagon, and the third value would be 1 because Enterprising and Realistic are two corners apart. These values are then used in the following formula: $C = 3(Value_1) + 2(Value_2) + (Value_3)$. Using the previous example, the corresponding formula would be $C = 3(3) + 2(2) + (1)$ resulting in a C index of 14. Scores can range from 0 to 18 with higher scores indicating greater interest-environment congruence. Holland (1997) endorsed the C index, stating that it "has a normal simulated distribution, is easy to calculate, is closely related to the theory, is sensitive to code orders, and has its highest correlations with the first letter on the hexagonal index" (p. 166).

Profile correlation. This congruence calculation method is the newest addition to the literature. Here, the Pearson correlation between an individual's RIASEC standard scores and a corresponding environmental profile score is calculated as the measure of congruence. Environmental profile scores are developed using a separate data set of individuals who have demonstrated success in the environment (i.e., incumbent method). In higher education literature, this has historically been done by researchers using data from the UNIACT. Here, profile scores for academic majors are calculated using a data set of students who have persisted into their sophomore year and have a minimum cumulative GPA of 2.00. Profile scores for all six RIASEC types are calculated for each major based on the mean scores of the students within them (Allen & Robbins, 2010). While "empirical reviews have pointed to the superiority of the profile correlation method," it is not possible in the absence of a separate data set that can be used to develop environmental profile scores (ACT, 2009).

Euclidean distance. Also seemingly used exclusively with UNIACT scores, this strategy for measuring congruence requires researchers to convert and plot individual and environmental profile scores as coordinate points on the data/idea and people/things dimensions. In the absence of environmental profile scores, estimates locations can be used by assigning them to ACT careers areas on the World-of-Work Map. Once plotted, the Euclidian distance is between these points is calculated to determine congruence. The smaller the distance, the greater the congruence. This method has also been criticized due to issues with interpretation (ACT, 2009).

There are a couple points worth noting before moving on to an overview of the application of Holland's theory to higher education. There are examples of other methods for measuring interest-major congruence (e.g., angular distance and composite scores), but these are typically one-offs that do not reappear in the methodologies of multiple studies (Allan & Robbins, 2008; Tracy & Robbins, 2006;). Profile correlation and Euclidean distance appear to have largely usurped traditional congruence indices as the prominent method for measuring congruence. However, these methods require access to large, separate datasets in order to develop environmental profile scores which is often not available. As a result, congruence indices maintain a valuable tool for researchers wanting to explore the applicability of Holland's theory at a smaller scale.

Holland and Higher Education

Higher education researchers have applied Holland's theory to understanding college student outcomes in the past. Feldman, Smart, and Ethington (2004) sought to explore the gains and costs in abilities and interests for both congruent and incongruent students with a sample of 2,309 students who had completed the 1986 and 1990 surveys of the Cooperative Institutional Research Program (CIRP). To identify gains, the authors used data from the CIRP to identify students' abilities, goals, and values that resembled the characteristics of the four RIASEC types (i.e., Investigative, Artistic, Social, and Enterprising) that were included in their analyses and coded them accordingly. Students were assigned personality types based on their highest score on the interest and ability scales. They found that, while congruent students did make gains in the abilities and interests associated with their major, incongruent students also made gains for their respective majors. This is consistent with previously conducted research by the authors and supports Holland's (1997) socialization assumption that environments reward and reinforce the characteristics associated with their respective RIASEC type.

Tracy and Robbins (2006) conducted a longitudinal study to look at the relationship between interest-major congruence, college achievement (cumulative GPA), and persistence with a sample 80,574 students enrolled in 87 bachelor's level degree granting institutions. The authors calculated congruence scores using the Euclidean distance method. In addition to this measurement of congruence, the authors also included two potential moderators of congruence-outcome relationship: interest profile clarity and interest level. Interest profile clarity was defined as the "magnitude of differentiation in the entire profile circular profile," and interest level the mean score across all six RIASEC scores (p. 67). Composite ACT score was also included in the analysis to determine if congruence could predict outcomes above what was predicted by this measure of academic achievement upon entering college. While the authors acknowledged that multiple or logistic regression is the typical analysis of choice for their type of research question, they chose to use hierarchical linear modeling due to the absence of independent data.

Tracy and Robbins (2006) conducted four hierarchical sets of analyses to identify the changes to the model after each predictor variable was added. Their results indicated that their measure of congruence added to the prediction of first-year GPA, second-year GPA, and GPA at graduation above what ACT score could alone. However, the same was not true for persistence as an outcome where "having a match between major and interests was not a predictor" (p. 85). Notably, the authors had included interaction terms and found that interest-major congruence mattered for individuals who had low overall interest levels.

Allen and Robbins (2008) investigated if interest-major fit related to third-year major persistence with a sample of 47,914students across 25 four-year institutions. Because students in their data set were nested within institutions, they used a hierarchical logistic regression model to account for institutional sources of variation. Rather than use a traditional measure of congruence, the authors calculated interest-major composite scores based on their entering major and work task scores. This interest-major composite was included alongside several other predictor variables including high school GPA, ACT composite score, major sureness, and first-year GPA to predict major persistence. The results showed that both the interest-major composite and first-year GPA were significant predictors of major persistence.

Allen and Robbins (2010) explored the relationship between interest-major congruence, motivation, first-year academic performance, and timely degree completion using longitudinal data from a sample of 3,072 students from 15 four-year institutions and 788 students from 13 two-year institutions. Motivation was measured using students' Academic Discipline score on the Student Readiness Inventory (SRI) which is calculated by summing the individual scores across 10 items designed to measure students' perceptions about the amount of effort they are putting into their academic work. Each student's RIASEC interest profile was obtained through their results on the UNIACT. Academic major, cumulative GPA, cumulative credit hours earned, and enrollment status were provided by the participating institutions in four waves. Interest-major congruence was calculated using profile correlation. Interest profiles for majors were determined using a separate dataset of 21,298 students from both four-year and two-year institutions.

Allen and Robbins (2010) used path models alongside hierarchical linear and logistic regression to evaluate the impact of these variables on first-year academic performance and timely degree completion. For both the four- and two-year samples, the overall models were significant, but interest-major congruence was not a significant predictor of first-year academic performance. Results were mixed for timely degree attainment with interest-major congruence being a significant predictor for four-year institutions but not two-year. It should also be noted that the overall models, while significant, explained far less of the variance for two-year institutions. Thus, outcomes were "more predictable" for the four-year sample. While the authors pointed out that this might be attributable to the smaller sample available for two-year institutions, they also acknowledged that “students who enter two-year institutions are more likely to have life situations that may affect postsecondary success that are not captured by our data set” (p. 33).

Tracey, Allen, and Robbins (2012) examined the relationship between interest-major congruence to indicators of college success with a sample of 88,813 undergraduates from 42 different colleges and universities. Both Euclidean distance and profile correlation were used as measures of congruence. GPA after one semester and two years, enrollment persistence after one and two years, and major persistence into year 3 were used as the indicators of college success. While the authors did not explicitly identify if these were all four-year institutions, these outcome variables suggest that they are. The authors included environmental constraint and individual flexibility as moderators. Environmental constraint refers to the extent that the environment is strong or weak in terms of how it impacts behavior. RIASEC environmental profile variance was used to determine environmental constraint with the authors asserting that majors with a consistent RIASEC profile (lower variance) are strong and inconsistent (high variance) as weak. Individual flexibility was operationalized as the overall, average level of a student’s interest score with higher averages indicating flexibility and lower averages not. Cumulative ACT score was also included in the analysis to determine what the variables could predict over and above this measure of college preparation.

Tracy, Allen, and Robbins (2012) employed hierarchical modeling to test their research questions. Each of the five outcome variables were assessed for both methods of P-E congruence (i.e., Euclidean distance and profile correlation). There was “mixed support” for the relation of congruence and GPA as evidenced by profile correlation being significant for first-semester GPA but not third year, and the reverse being true for Euclidean distance. While both congruence measures were significant for major persistence into the third year, neither were significant for both second year and third year enrollment persistence. However, the interaction term for congruence by environmental constraint was significant for both measures of congruence across all three persistence outcomes. This supported the hypothesis that congruence matters more for majors had greater environmental constraint. They also found that congruence was more significant for less flexibility (i.e., lower profile levels).

Milsom and Coughlin (2017) looked at the relationships between student-major congruence, consistency, differentiation, academic major satisfaction, and academic performance with a sample of 99 undergraduate students at a large southeastern university. Academic major satisfaction was measured using the Academic Major Satisfaction Scale (AMSS). The AMSS (Nauta, 2007) measures global satisfaction with one's major using six, 5-point Likert scale items ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). Consistency was calculated using the Hirschi and Läge's (2007) method. Here, a value from 1 to 3 is assigned based on the hexagonal relationship between the top two RIASEC types in an individual's profile. Adjacent types are assigned a 3, alternate types a 2, and opposite types a 1. Differentiation was calculated using the Streuungs Index. In this method, the standard deviations for the RIASEC types are calculated with higher values indicating greater differentiation. The SDS (Holland, 1994) was used to generate RIASEC scores for each participant, and the C index (Brown & Gore, 1994) was used to calculate a congruence score with student's major. Major codes were determined based on *The Educational Opportunities Finder* (Rosen, Holmberg, & Holland, 1997 as cited in Milsom & Coughlin, 2017).

Milsom and Coughlin (2017) utilized two separate linear regressions. In the first, congruence, consistency, and differentiation were used as predictors for AMSS scores, and in the second, congruence, consistency, differentiation, and AMSS scores were used to predict academic performance, measured as self-reported, cumulative GPA. The first analysis was not significant; however, the second analysis found that both congruence and AMSS scores were significant predictors of GPA. The model accounted for 12% of the variation in GPA, $R^2 = .12$, $F(4, 90) = 2.27$, $p = .02$. Neither consistency nor differentiation were significant in either regression.

Rocconi, Liu, and Pike (2020) conducted a study to explore the relationship between perceived learning gains, satisfaction with college, students' grades, college experiences, person-major congruence, and student characteristics. Rocconi, Liu, and Pike (2020) conducted a study to explore the relationship between perceived learning gains, satisfaction with college, students' grades, college experiences, person-major congruence, and student characteristics. Using data from the 2015 National Survey of Student Engagement (NSSE), the authors put together a sample of 3350 senior students from four-year colleges and universities. This survey had included a Holland item set which allowed the authors to assign each student a RIASEC profile. Reported majors were assigned a two-digit Holland code based on the Dictionary of Holland Occupational Codes. Congruence was calculated using Wiggin and Moody's (1981 as cited in Rocconi, Liu, and Pike, 2020) congruency index. This index uses the dominant and secondary RIASEC types for both the individual and the environment to assign the individual a score of 0-5 to represent their level of congruence with 0 indicating no congruence and 5 perfect congruence.

For their analysis, the Rocconi, Liu, and Pike (2020) used a fixed effects regression model. GPA was estimated using students' self-reported grades. A perceived learning gains score was determined using the sum of students' ratings on ten questions related to the topic. Similarly, the authors used the sum of student's rating on two questions dealing with their overall educational experience and returning to the same institution to calculate a score for satisfaction with college for each student. Congruence was found to have a significant relationship with grades but not perceived gains in learning or satisfaction.

Several notable themes and takeaways emerged from this review of studies that have applied Holland's theory to the study of students in higher education. In general, there seems to be mixed support for Holland's theory when it comes to predicting higher education outcomes. The socialization assumption seems to certainly apply to college majors as both congruent and incongruent students have been shown to make gains related to their major over time (Feldman, Smart, and Ethington, 2004). Furthermore, interest-major congruence is consistently a significant predictor of major persistence (Allen and Robbins, 2008; Tracy, Allen, and Robbins, 2012). However, the verdict is still out for the outcomes that are often most important to higher education researchers: academic performance and enrollment persistence.

While the majority of studies found congruence to have a significant relationship with GPA (Allen, & Robbins, 2012; Rocconi, Liu, & Pike, 2020; Tracy, Milsom & Coughlin, 2017; Tracy & Robbins, 2006), there is at least one example of it not (Allen & Robbins, 2010). Likewise, some studies showed that congruence had a significant relationship with enrollment persistence (Allen & Robbins, 2010) while others found no such relationship (Tracy, Allen, & Robbins, 2012; Tracy & Robbins, 2006). Thus, there is still a great opportunity to contribute to the literature on congruence and higher education outcomes to further illuminate the uncertainty that there seems too currently be.

Furthermore, most of these studies were completed exclusively with four-year institutions (Allen & Robbins, 2008; Milson & Coughlin, 2017; Rocconi, Liu, & Pike, 2020; Tracey, Allen, and Robbins, 2012; Tracy & Robbins, 2006). The one study that included a two-year student sample found both that interest-major congruence was not a significant predictor of persistence like it was for four-year institutions and that their overall models for both GPA and persistence explained far less of the variance for two-year institutions than four-year (Allen & Robbins, 2010). All of these facts reinforce the statement of the problem articulated in the opening chapter. Not enough research has been done with PTYCs, and the little bit that has been done affirms that the academic outcomes of students at these institutions cannot be predicted as reliably using models derived from four-year institutions. This appears to also be true for the application of person-academic environment congruence.

Finally, every reviewed study shared one methodological characteristic, defining the academic environment as the student's chosen major. This presents its own set of issues that will be discussed in greater detail in the next section.

Defining and Classifying the Academic Environment

Whereas a career represents a relatively stable environment with a consistent set of activities, tasks, and roles, the academic environment is far more diverse. For example, it would be foolish to expect an entire institution to have a well differentiated Holland code due to the range of curricular and cocurricular activities, tasks, and roles that are available. As highlighted in the closing paragraph on classifying environments, Holland (1997) recognized that many environments are diverse in character and provided recommendations for helping researchers and practitioners when attempting to apply the concept of congruence. Most relevant is the importance of paying attention to where an individual is in these large, heterogenous environments and the experiences that they are having as a result. These subenvironments represent "the largest or most influential portion of a person's environment" (Holland, 1997, p. 42). As the previously reviewed studies demonstrate, research on congruence in higher education has traditionally attempted to account for subenvironments by defining the student's academic environment as being their identified instructional program. Nye, Su, Rounds, and Drasgow (2012) sum up this approach when they state that "college majors provide a salient and proximal environment to the person" (p. 385).

However, defining the student's academic environment as their instructional program of choice has multiple drawbacks and methodological issues. First, requiring an academic major means that undecided students must be excluded from studies. When comparing their final sample to the students that they had to excluded due to either undecided or dropout, Allen and Robbins (2008) found that undecided students were more likely to be minority students, have a lower high school GPA, lower ACT score, and lower first-year GPA. Thus, undecided students appear to be a more vulnerable population. A methodological approach that excludes them by default not only limits the generalizability of any results but also misses out on the opportunity to explore the factors that influence persistence for students for whom there is a particularly strong need to better understand.

Furthermore, there is skepticism in the literature that the major accurately represents the academic environment for all college students. While this makes conceptual sense, it has limitations, especially when it comes to community college students. As Tracy and Robbins (2006) point out, "college major choice is not the best representation of occupational environment" because the "college environment has many opportunities to interact in a wide variety of environments and the major is not as constraining as a job in that area would be." In other words, a student's major is not always indicative of the *actual* environment that they are in.

This is especially true in the first year due to general education requirements. As Allen and Robbin (2008) point out, students "typically take several general education courses during their first year, so *first-year GPA is often a measure of performance in general courses rather than performance in major* [emphasis added]" (p 76). PTYC students, who tend to be more academically unprepared, often have to spend more time in these environments due to the need for remedial education (Skomsvold, 2014). This means that, even if a student does have a chosen instructional program, it is far from an accurate representation of their actual academic environment if they are primarily enrolled in remedial and general education courses their first year. Thus, a new approach to defining and classifying the academic environment is needed that is inclusive of undecided students and accounts for the heterogeneous nature of the first year.

If an instructional program is a subenvironment of the institution, then classes are even more precise subenvironments of the instructional program. It has been found that "classroom environments reflect systematic differences among various field of study" (Astin, 1965, p. 280). Thus, one should be able to categorize a specific class based on the larger instructional program that it is *associated* with the same way that these instructional programs were originally categorized based on their associations with certain occupations. One might argue that an incumbent method would call this into question, especially in a general education class where the largest number of members are students who likely represent a diverse range of interests. However, as pointed out earlier, Holland (1997) was adamant that one must pay attention to the "distribution of power" when categorizing an environment. He stated that environments are "defined as the situation or atmosphere created by the people who dominate" them (p. 41). A classroom is traditionally composed of a professor and students, and even though these students far outnumber the professor, it is the professor that "dominates" that environment through their position of authority. Through this authority, the professor defines that environment by requiring the students to complete the requirements of the course, which reflect the major and, by association, a range of occupations.

"Other Things Being Equal…"

The title of this section is a direct reference to a phrase that Holland (1997) uses multiple times throughout his seminal book. It serves as both a preface to many of his assertions about the theory and as a section title as well. It refers to the acknowledgment that a variety of other individual characteristics influence behavior in environments and that these factors must be taken into consideration when applying the theory and the concept of congruence both in practice and research. While not exhaustive, this review provides an overview of research that has been done on inequalities in educational outcomes among various subgroups at community colleges.

A quick overview of salient statistics demonstrates that there are numerous observable inequalities when outcomes are disaggregated by gender, age, and level of academic preparation. Women graduate at a much higher rate when compared to men as evidenced by that fact that, during the 2009-10 academic year, women accounted 62% of all associates degrees earned (Cohen, Brawer, & Kisker, 2013). "Traditional" students, defined as students who enroll the fall semester following their graduation from high school, have much higher associate degree (18.4%) completion rates when compared to "adult" students (10.9%), defined as students who enrolled at 24 years or older (Hoachlander, Sikora, & Horn, 2003). Academic preparation has long been understood to be a strong predictor for college outcomes and many underprepared students enroll at PTYCs. For the 2011-12 academic year, approximately 41% of all community college students took a remedial course (Skomsvold, 2014). Research has shown that only 9.5% of community college students who enroll in remediation coursework go on to graduate compared to 13.9% of college ready students (Adams et al., 2012).

There are also disparities in first-year retention among community college students based on race/ethnicity. For the Fall 2016 cohort, Asian (57.3%) and Hispanic (53%) students had the highest retention rate followed by white (50.2%) and black (42.1%) students (NSC Research Center, 2018). Looking at degree attainment paints a somewhat different picture with white students earning a disproportionately higher percentage of associate degrees compared to their percentage of total enrollment than other ethnicities (Cohen, Brawer, Florence, Kisker, 2013). First-generation college students (FGS) also have concerning outcomes when compared to their continuing-generation peers. Among college students who began enrollment in the 2003-04 academic year, a higher percentage of FGS (33%) had dropped out within three years compared to continuing-generation students (CGS) with parents who had same college education (26%) or had earned a bachelor's degree (14%) (Cataldi, Bennet, & Chen, 2018).

Interestingly, SES appears to play less of a role in degree attainment at the associate degree level. The Education Longitudinal Study of 2002 (ELS:2002) followed a nationally representative cohort of 10th grade students from 2002-2012. Bachelor's degree attainment did vary tremendously based on level of SES with approximately 60% of high SES students completing a bachelor's degree or higher by 2012 compared to only 29% of middle SES and 14% of low SES students. However, the differences were far less pronounced in terms of associate's degree attainment with low (8%), middle (10%), and high (7%) SES students all having low, but comparable rates (NCES, 2015). Bailey, Jenkins, and Leinbach (2005) found similar results in their review of student outcomes among students who began their higher education career at a community college. They found that the percentage of students whose highest outcome attained in six years was an associate degree was fairly consistent across the lowest (19%), second (13%), third (18%), and highest (14%) household income quartiles and concluded that “there does not seem to be a clear correlation between household income and the rates at which students earn associate degree” among community college students (p. 46).

With the exception of SES, the literature shows that each of the groups reviewed have inequalities in academic success at community colleges. Female students are more successful than male students, traditional students tend to more successful than post-traditional students, academically unprepared students are less successful than prepared, white students account for a disproportionately higher percentage of associate degrees earned, and FGS students drop out at higher rates when compared to CGS.

Holland (1997) specifically cites Gottfredson (1981) and her concept of “circumscription” in his discussion of how individual characteristics can limit the careers or, when translated to the topic of this study, academic environments that an individual is willing to consider as a viable option. Gottfredson asserts that, as an individual’s self-concept develops and becomes more complex, new criterion are adding for judging whether an occupation is compatible with their view of self. The consequence is that the range of occupations that are acceptable to that individual becomes narrower over time. As Gottfredson and Lapan (1997) put it, circumscription is the process of the individual determining what is “socially unacceptable for people like them” (p. 425). For example, the first stage of this process is when children stop considering “magical” jobs and limit their aspirations to traditional “adult” occupations. Gender, perceived prestige, difficulty, and, finally, interests and abilities, continue to limit acceptable occupations throughout an individual’s development (Holland, 1997).

Gottfredson and Lapan (1997) assert that circumscription by gender results in substantially different range of interests between men and women due to gender-based socialization. Evidence of this can be easily observed in major distributions among men and women at community colleges. In the 2009-10 academic year, women obtained a disproportionate number of the degrees in education (86%), health professions (85%), and business (65%). Similar outcomes can be observed for men in fields like mechanic and repair technologies (95%), engineering technology (87%), computer and information science (76%), and physical sciences (62%) (Cohen, Brawer, Florence, Kisker, 2013). Hacket & Betz (1981) proposed that gender-based socialization experiences result in a lack of *self-efficacy* in career-related behaviors that are not considered culturally appropriate for women. Self-efficacy and its relationship to culture has received a great deal of attention in the literature.

Self-efficacy refers to an individual's belief or beliefs about their ability to successfully complete tasks. It is theorized that higher levels of self-efficacy result in higher levels of achievement (Bandura, 1997; Lent, Brown, & Hackett, 1994). When applied to higher education, researchers explore college students' self-efficacy beliefs concerning their ability to be successful in the college environment. Self-efficacy has been showed to be significantly related to both academic performance and persistence (Multon, Brown, & Lent, 1991). Furthermore, self-efficacy beliefs have been shown to vary based on culture.

There are multiple examples of this in the literature. Underprepared students rate themselves significantly lower in areas such as academic ability, intellectual self-confidence, and achievement drive when compared to college-ready students (Grimes & David, 1999). Nonminority undergraduate women often anticipate lower scores on reading comprehension and creativity tasks compared to other nonminority women, nonminority men, and both minority men and women. Similarly, both undergraduate minority men and women expected to score significantly lower compared to other groups, but also predicted poor performance for all individuals in their respective groups. In contrast, nonminority men predicted they would perform as well as others in their group and assumed that this level of performance was typical for the average undergraduate (Mayo & Christenfield, 1999). Finally, FGS, community college students have been shown to have lower self-efficacy when compared to their peers (Inman & Mayes, 1999).

It should be noted that, with the exception of Inman and Mayes (1999), the previously cited literature was conducted with students in four-year institutional settings. There have been recent attempts at exploring the relationship between psychological factors, including self-efficacy, and community college retention. Luke, Redekop, and Burgin (2014) collected survey data from 1,191 students at a Mid-Atlantic community college that included a measure for career decision self-efficacy (CDSE). Student who reported higher CDSE were *less likely* to report that they intended to return to the institution the following semester. The authors hypothesized that this might reflect that either students felt confident enough to not need the degree or that it indicated unrealistic beliefs about their own abilities. Furthermore, they found no significant differences in CDSE based on age, gender, ethnicity, or parental education.

Researchers have also applied measures of *academic* self-efficacy to understanding community college persistence. Similar to Luke et al. (2014), Nakajima, Dembo, & Mossler (2012), found the academic self-efficacy was not a significant predictor of persistence with a sample of 427 students from a single community college in California. However, they did observe that self-efficacy was significantly related to cumulative GPA, which was the strongest predictor of persistence in their study. Thus, they hypothesized that the relationship between academic self-efficacy and persistence is mediated by cumulative GPA.

Evidence of the impact that culture can have on choice of preferred environment can be found in other research methodologies that do not utilize Gottfredson’s concepts or self-efficacy as a theoretical framework. Porter and Umbach (2006) looked at the relationship between college major choice and variety of variables, including gender, ethnicity, academic preparation, and RIASEC types. They found that, after controlling for demographics, family influence, academic self-efficacy, political views, and the RIASEC types, the initially significant differences in major selection based on gender and academic preparation disappeared. However, differences among ethnic groups remained significant even after including the previously mentioned variables. Black students were more likely to choose interdisciplinary and social science majors over science majors when compared to white students, and Hispanic students were more likely to choose arts & humanities, interdisciplinary, or social science majors than science majors when compared to white students.

Individual differences are not always psychological. Concrete obstacles and the pragmatic considerations that come with them can be cultural. For example, FGS students report different reasons for dropping out of college at higher rates when compared to CGS. Approximately 54% of FGS indicated that they "Couldn't afford to continue going to school" compared to 45% of CGS. "Change in family status," which includes marriage, having a baby, or a death in the family, is also endorsed at a much higher rate by FGS (42%) than CGS (32%). Similarly, Terenzini, Springer, Yaeger, Pascarella, and Nora (1996) found that FGS tend to have more dependent children and work more hours per week off-campus.

In addition to these pragmatic concerns, FGS students also lack the "cultural capital" that CGS students have when it comes to knowing how to navigate the world of higher education. Cultural capital is defined as a "general familiarity with the traditions and norms necessary to be successful" (Mehta, Newbold, & O'Rourke, 2011, p. 21). In higher education, cultural capital can include general knowledge about financial aid, registering for classes, and availability of student services as well as effective strategies for coping with stress. Furthermore, because FGS have more external demands on their time (e.g., working, family, etc.), they are less involved on campus, inhibiting their ability to gain the capital that they were not able to receive from their parents. Additionally, FGS students perceive that their families are less supportive and encouraging of academic supports when compared CGS (Mehta, Newbold, & O'Rourke, 2011). Research has also shown that African Americans and Hispanic students have social capital disadvantages when compared to their white peers and that this can negatively impact college enrollment due to a limited access to valued resources (Perna & Titus, 2005). This is unsurprising given that FGS have an increased likelihood of representing an ethnic minority (Bui, 2002).

The numbers and research affirm Holland's (1997) point that cultural influence and individual characteristics should be identified and accounted for. He encouraged researchers to include measures of these cultural characteristics in order study their influence in a more "integrated way" (p. 13-14). When applied to a study, researchers should control for these variables to determine what relationship congruence has with academic outcomes.

Conclusion

The goal of this chapter was to provide an overview of the relevant literature on community college student persistence and establish why a P-E fit theory, Holland's (1997) in particular, could provide a novel theoretical framework for researching this topic. An overview of methods for classifying individuals, environments, and the congruence between was provided, highlighting strengths and weaknesses. Higher education researchers have applied Holland's concepts to their topics in the field. These have proven to be insightful but have also demonstrated a need to reconsider how the academic environment is conceptualized. Finally, Holland emphasized the need to take into consideration the impact of culture and other relevant characteristics when using this theory. The community college literature has confirmed that this well advised and that cultural and precollege characteristics should be taken into account and controlled for when attempting to look at the relationship between another independent variable and academic outcomes.

CHAPTER 3

RESEARCH METHODS

Introduction

The purpose of this study was to better understand the contributing factors to first-year student performance and stability at a large, public community college by analyzing the relationship between John Holland's (1997) concept of congruence and these outcomes after controlling for cultural and precollege characteristics. This study used a quantitative design and be guided by two research questions:

1. Does congruence between students' occupational personality type and their first-semester academic environment predict first-semester GPA?
2. Does congruence between students' occupational personality type and their first-semester academic environment predict persistence into the second year?

This chapter will provide a detailed overview of the data, methods and procedures that will be used in conducting this study including the sources of data, sample, research design, data analysis, and limitations.

Site

The site for this study was Pellissippi State Community College (PSCC). Per the Carnegie Classification of Institutions, PSCC is a large, public two-year institution located in the southeast region of the United States.

Data Source

The opportunity to perform this study was the result of PSCC's participation in Complete College America's (CCA) Purpose First initiative. CCA was established in 2009 with the goal of "increasing college completion rates and closing equity gaps by working with states, systems, institutions, and partners to scale highly effective structural reforms and promote policies that improve student success" (Complete College America, n.d.). The Purpose First initiative sought to identify and implement strategies for empowering students to make informed decisions about their academic and career goals earlier on in their college experience (*Purpose First*, 2020). To address the goals of this initiative, PSCC asked students to complete the Purposeful Advising With Students (PAWS) module as part of their admissions process. PAWS was designed to encourage and help new and returning students engage in career exploration early in their college career. It included a short video designed to review important career decision-making concepts. Students were then directed to complete a version of the O*Net Interest Profiler Short Form (IP Short Form). After finishing the assessment, students were asked to input their top three RIASEC types from highest to lowest using three dropdown boxes based on their IP Short Form results. Completing this survey marked the student as having completed the module. This information was stored in a dataset along with the student's name and ID number and the date that they completed PAWS. However, students' actual results were also stored for review on a one-by-one basis.

The Institutional Review Board process was then completed at the site of the study, Pellissippi State Community College. Upon receiving the approval letter (see Appendix B), this list of students was sent to Pellissippi State's Institutional Effectiveness, Assessment, and Planning (IEAP) Office to obtain demographic, academic outcome, and course enrollment data. The IEAP returned this data with identifying information removed.

Instrumentation

The IP Short Form was developed in 2010 and contains 60 items, 10 for each of the six RIASEC types. Each item describes a work activity, and participants rate their interest on a five-point scale: (1) Strongly Dislike, (2) Dislike, (3) Unsure, (4) Like, and (5) Strongly Like. These responses are then converted into numerical values from 1 (Strongly Dislike) to 5 (Strongly Like). Points are summed up for each RIASEC category. Thus, the score for each category ranges from 10 to 50 based on how strongly the individual preferred the activities associated with it. Categories are then arranged from the highest score to the lowest score to create their unique profile (Rounds et al, 2010). The paper-and-pencil version of the IP Short Form can be found in Appendix A.

Sample

The initial sample included 1,604 individuals who had completed PAWS between the date it launched, 6/7/2018, and the day before classes began for the Fall 2018 semester, 8/26/2018. This list of students was provided to the college's institutional research office who then provided the additional variables including ethnicity, first-generation status, Pell grant eligibility, high school GPA, first semester GPA, and Fall 2019 enrollment status. Not every individual who completed PAWS would go on to take courses in the Fall 2018 semester. Thus, these students were removed from the sample leaving 1,406 students with Holland Code data. However, there were several concerns about the data as at this point.

First, recall that "completing" PAWS simply means that the student filled out the survey at the end of the module. The accuracy of this survey information was dependent on the student taking the full assessment, returning to the module, and correctly selecting the top three traits in the correct order based on their results. This process invites problems that can result in inaccurate data. At best, well intentioned students may input the wrong results due to user error, and at worst, savvy students seeking to quickly complete PAWS may randomly pick their RIASEC types after taking the IP Short Form or skipping it altogether. To further complicate things, the menu that students use to select and begin the IP Short Form also includes the option to take a much shorter, 6-question assessment that also provides them with a Holland Code. Explicit instructions were included in PAWS to not take this assessment; however, there was still concern that students would complete it instead and input their RIASEC types based on these far less reliable results. For these reasons, extra steps were taken to clean and ensure the accuracy of the data.

Each of the remaining 1,406 students were manually looked up to determine if they had taken the correct assessment and what their Holland Code was. This made it possible to not only record their top three RIASEC types, but also their complete profile with scores for each type. These scores ranged from 0-100. Students with no identifiable account were removed from the sample. Furthermore, students who had an identifiable account but had either completed the 6-question assessment or had no recorded results were also identified and removed from the sample. This resulted in 1,257 students with complete and accurate Holland Code data.

Table 3.1 provides descriptives for the categorical independent variables to be used in the study. The sample was predominantly female (57.9%), white (80.5%), first-generation (62.3%), and in the traditional age range (82.2%). Per the institution's electronic fact book, this is fairly consistent with demographics for the overall student body during the Fall 2018 semester which was also largely female (51.7%), white (79.9%), first-generation (54%), and in the traditional age range (70.6%). Since the sample was composed of students who had completed PAWS as part of their new student orientation process, it was made up of a mix of new, returning, and transfer students and few to no continuing students. Thus, deviations from the overall student body are likely attributable to unequal persistence among these groups. For example, students within the traditional age range made up 82.2% of the sample compared to only 70.6% of all students. A

Table 3.1
Descriptives for Categorical Variables

Variable	Label	Count	%
Gender	Female	617	57.9%
	Male	448	42.1%
Ethnicity	American Indian/Alaskan Native	2	0.2%
	Asian	23	1.8%
	Black/African American	83	6.6%
	Hispanic	97	7.7%
	Native Hawaiian/Other Pacific Islander	1	0.1%
	White	1004	79.9%
	Multiracial	46	3.7%
First-Gen Status	First-Generation	664	62.3%
	Continuing-Generation	401	37.7%
Age	Traditional Student	875	82.2%
	Post-Traditional Student	180	17.8%

review of persistence data within this same fact book shows that post-traditional students persist at higher rates when compared to traditional students. Thus, they would make up a greater percentage of the total student body over time, which is consistent with the observed data.

Table 3.2 shows the breakdown of dominant, secondary, and tertiary types for the sample. Reviewing this information suggests that students, based on RIASEC type, do not appear to enroll in community colleges at equivalent rates. Interestingly, the most represented types of students, from largest to smallest, were Social (28.4%), Investigative (22.12%), and Artistic (18.38%), and the most underrepresented was Realistic (6.6%).

As detailed in Chapter Three, an overall congruence score was calculated for each student for the Fall 2018 semester by calculating the C Index scores between each student and each individual course that were enrolled in, multiplying each score by the credit hour value of the respective course, summing these adjusted scores, and dividing it by the total number of credit

Table 3.2
RIASEC Descriptives

	Dominant		Secondary		Tertiary	
	Count	**%**	**Count**	**%**	**Count**	**%**
Realistic	83	6.60%	132	10.50%	183	14.56%
Investigative	278	22.12%	219	17.42%	173	13.76%
Artistic	231	18.38%	194	15.43%	187	14.88%
Social	357	28.40%	293	23.31%	223	17.74%
Enterprising	159	12.65%	249	19.81%	303	24.11%
Conventional	149	11.85%	170	13.52%	188	14.96%

hours that the student is enrolled in to arrive at an overall congruence score for the semester. The following formula represents how to calculate this overall congruence score.

Overall Congruence Score (c) = $(a_1 + a_2 \ldots a_k)/h$

a = adjusted score

$_k$ = the number of adjusted scores

h = total number of credit hours

It should be noted that there were many cases where there were ties between RIASEC types, resulting in an arguably arbitrary order.

Table 3.3 shows the descriptives for the continuous independent variable, congruence score. The mean for the congruence score was 9.9 (SD = 2.5). An initial review of descriptives for the high school GPA variable revealed values that were not consistent with the 4-point scale (e.g., 670). A descriptives table for this variable will be generated after it has been reviewed in the data cleaning stage of the analysis in Chapter Four.

Table 3.3
Descriptives for Congruence Score

	N	Mean	SD	Range
Congruence Score	1257	9.9	2.5	2-18

Variables

Dependent Variables

High congruence between an individual and their environment is theorized to result in several desirable outcomes including increased performance, and stability (Nauta, 2013). In an academic environment, students' grades are the best available indicators of their performance, and their grade-point average (GPA) provides a quantitative measure of their overall performance in that academic environment. Similarly, student stability at a college or university is traditionally referred to as persistence. Pascarella and Terenzini (2005) defined persistence as the "progressive reenrollment in college" (p. 374). Thus, first-semester GPA and persistence from the first semester to the second academic year were chosen as the two dependent variables.

The first semester was chosen because it has been shown to be "especially important in terms of student adjustment and ultimately fostering student success" (Bowman, Jarratt, Jang, & Bono, 2019, p. 274). First-semester GPA has been shown to be a statistically significant predictor of 6-year graduation rate for underrepresented students (Gershenfield, Ward Hood, & Zhan, 2016). Furthermore, the need for consistent environmental data (i.e., course enrollment) necessitates choosing a single semester rather than the whole academic year. Using the whole academic year would require excluding students who did not persist into the second semester as they would have incomplete environmental data. Persistence into the second year was chosen over persistence into the second semester because it is a more valuable and accurate representation of a student's long-term academic stability. Also, the second year is when many students first begin to experience the negative consequences of poor academic performance (i.e., academic and/or financial aid suspension) that prevent them from continuing even if the desire to do so remains.

Independent Variables

Cultural and precollege characteristics. As the literature review in Chapter Two showed, there is a relationship between cultural and precollege characteristics and student success outcomes. Furthermore, the theoretical framework calls for consideration of these types of variables when applying the theory. Holland (1997) listed age, gender, ethnicity, geography, social class, physical assets/liabilities, educational level attained, intelligence, and influence as variables that require direct assessment in order properly apply the theory and study their influence in a more "integrated way" (p. 13-14). Of these, several will be excluded from consideration from the study due to being either not pertinent or currently inaccessible through available data. However, age, gender, and ethnicity will be included along with an indirect measure of social class. Furthermore, parent-education level and high school GPA will also be included due to their relevance in higher education.

Age will be represented by a dichotomous, categorical variable indicating if the student is a traditional (23 or younger) or post-traditional (24 or older) learner. These categories align with the two "free college" initiatives that Tennessee has implemented over the last several years: (1) Tennessee Promise which is aimed at traditional students coming straight out of high school and (2) Tennessee Reconnect which is aimed at post-traditional students, defined as 24 years or older, who do not have a college credential. Thus, there is a policy-driven need to look at the outcomes for these distinct groups of students and this categorization makes comparative analyses easier.

Gender will also be a dichotomous, categorical variable based on if the student has identified as male or female. Ethnicity is a categorical variable that includes white, black or African American, Hispanic or Latino, Asian, American Indian, and Multiracial. The best available indicator of social class is if a student has ever been eligible for a Pell Grant. The Tennessee Higher Education Commission (THEC) uses “ever Pell-eligible” to categorize a student as low-income or not (*Tennessee Higher Education Fact Book*, 2020, p. 94). This study will do the same. Students will be categorized as either a first-generation or continuing-generation for parent education level. Finally, high school GPA will be included as a continuous variable.

Congruence. Holland (1997) asserted that individuals and environments can be described in terms of their resemblance to each of six model types: (1) Realistic, (2) Investigative, (3) Artistic, (4) Social, (5) Enterprising, and (6) Conventional. Together, these six types are typically referred to as the RIASEC types. When used to describe personality, the types refer to an individual’s attitudes, interests, and competencies. When used to describe environment, the types refer to the activities, tasks, and roles that are required. Congruence is then determined by comparing an individual’s RIASEC profile to that of the environment.

Calculating congruence between an individual and their environment requires that both be accurately defined and categorized using the RIASEC types. Students will be categorized using the Holland typology based on their results from taking the O*Net Interest Profiler. Each student will be assigned a Holland code based on three highest scored RIASEC types.

Environment can be immediately defined as the student’s *academic* environment. However, this, in and of itself, is not enough to provide clarity. This is in large part due to the heterogeneity of the college experience. Diversity is a value of higher education and is demonstrated through the wide range of opportunities that are available to the student both inside and outside of the classroom. While the literature has traditionally used a student's major as their academic environment, this has been shown to have methodological issues for students who are in their first year or have an undeclared major. Based on Holland's (1997) recommendation to assess for the precise subenvironments that represent the "most influential portion of a person's environment," the academic environment will be further narrowed to focus exclusively on the coursework that the student is taking. Classes will be assigned a Holland code based on the larger instructional program that is associated with. Holland Codes will be assigned to instructional programs using either the *Dictionary of Holland Codes* (DHOC), the *Educational Opportunities Finder* (EOF), or classification equations based on the same association method used in the DHOC. For each major, the DHOC was referred to first, followed by the EOF, and if neither resource had an adequate listing, the crosswalk between the most recent revision of the CIP and O*Net job analysis data was used to classify that major based on its associated occupations. See Appendix B for a summary table of PSCC's majors and their corresponding Holland Codes.

With how this study will define "person" and "environment" established, it is now possible to outline how fit, or congruence, will be calculated. Based on the literature review, the C index emerged as the most supported formula for calculating the congruence between the Holland codes of a person and an environment. However, due to the heterogeneity of the coursework of community college students, a single calculation of congruence is not possible. Instead, this study will use the C index to calculate the congruence between a student and each of their courses, resulting in separate scores. Based on Holland's (1997) recommendation to account for "time spent" in an environment, each of these scores would then be multiplied by the number of credit hours they are worth in order to account for the different time commitments for each. For example, two different courses could result in the same initial score of 13 for a particular student. However, if one course is worth 3 credit hours and the second is worth 1 credit hour, then it makes little sense to weight them equally in calculating the overall score for the student's academic environment. Thus, the initial score for the 3-credit hour course would be multiplied by 3 to result in a score of 39, and the initial score for the one-hour course would be multiplied by 1 to stay at 13. These various scores would then be averaged together to create an overall measure of congruence between the student and their academic environment that semester. This overall measure can then be used to evaluate the relationship between congruence and indicators of academic success.

Consider a student with the Holland code IEC and who is enrolled in English Composition I, Introduction to Political Science, and Survey of College Mathematics. Per the DHOC, the corresponding subjects for each of these courses yield the following Holland codes.

English: AES

Political Science: SEI

Mathematics: IRE

Figure 3.1 shows the calculation of the student's congruence score for each course using the C Index. The overall congruence score would then be calculated using the following formula.

Overall Congruence Score (c) = $(a_1 + a_2 \dots a_k)/h$

Here, the variables represent the following.

a = adjusted score

$_k$ = the number of adjusted scores

h = total number of credit hour

Subject	Hours	Subject Code	Student Code	C Index Calculation	C Index * Hours	Adjusted Score (a)
ENGL	3	AES	IEC	3(2) + 2(3) + 1 = 13	13 * 3	39
POLS	1	SEI	IEC	3(1) + 2(3) + 1 = 10	10 * 1	10
MATH	3	IRE	IEC	3(3) + 2(1) + 2 = 13	13 * 3	39

Figure 3.1
Example Student Congruence Calculation

For this example, the formula would look like the following.

Overall Congruence Score (c) = (39 + 10 + 39)/7

This would result in an overall congruence score of $c = 12.57$.

Data Cleaning and Preliminary Analysis

Data will be checked and cleaned within SPSS (version 24) using a combination of Tabachnick and Fidell's (2018), Keith's (2015), and Morrow and Skolits' (2017) guidelines. Frequencies will be generated for each variable to check for coding issues, missing data, and any other initial errors. Variables will be modified or recoded into new variables as needed to address these problems.

If data is missing from one or more continuous variables, Tabachnick and Fidell (2018) recommend using the Missing Variable Analysis (MVA) tool in SPSS to determine how pervasive and random the missing data is. If data is missing from a small percentage of the sample and appears missing completely at random (MCAR) or missing at random (MAR), listwise deletion of cases will likely be the most obvious choice. In listwise deletion, cases with missing data for any of the variables being used in the analysis are removed (or deleted) (Keith, 2015). A benchmark of greater than 5% will be used to determine if the amount of missing data is too significant for listwise deletion. Tabachnick and Fidell (2018) state that, when there is 5%

or less missing data, "almost any procedure for handling missing values yields similar results" (p. 54).

However, if the cases with missing data exceed 5% or appear to be missing not at random (MNAR), data estimation will be used. Other available demographic data (e.g., ethnicity, age, etc.) will be used to determine if there are patterns among the missing data. Additional research and consultation with the institution's institutional research office will be done as needed. If a pattern is able to be identified, group means will be developed and imputed for the missing values. If not, the grand mean will be used. Variables with missing categorical data will be removed using listwise deletion (Morrow & Skolits, 2017).

Descriptives and standardized scores will then be used to determine if there are any univariate outliers among the continuous variables. Tabachnick and Fidell (2018) state that cases "with standardized scores in excess of 3.29… are potential outliers" (p. 64). If outliers are identified, these values will be modified to be one "unit" larger or smaller than 3.29 standard deviations from the mean. Per Tabachnick and Fidell (2018), this method allows for the scores to remain "deviant, but not as deviant as they were" (p. 67).

The multivariate normality, linearity, and homoscedasticity assumptions will be checked using visual inspection of both the scatterplot and histogram of the residuals. A preliminary multiple regression will be conducted to generate these figures. If the points on the scatterplot are rectangularly shaped and concentrated along the center and the histogram is approximately bell-shaped and symmetrical, these assumptions are met (Tabachnick and Fidell, 2018). However, if the shape or pattern of either the scatterplot or histogram of residuals suggest a deviation from one of these assumptions, an analysis of the individual variables will be conducted.

When analyzing individual variables, histograms will be generated and inspected to identify problems. If the distribution of a variable appears to be significantly skewed, more precise measures of will be used to confirm if this is the case and how extreme the issue is. Tabachnick and Fidell (2018) recommend using significance tests for both skewness and kurtosis. A z-score for each statistic can be calculated by dividing the absolute value of the skew and kurtosis statistics by their corresponding standard errors and comparing that to an established cut-off point. Mayer (2013) asserts that 3.29 is an appropriate cut-off for sample sizes greater than 100. If the z-scores for skewness and/or kurtosis exceed this cut-off point, univariate normality will be considered not met. Keith (2015) highlights three methods for dealing with excessive nonnormal data: (1) transforming the dependent variable, (2) eliminating subgroups, or (3) using an alternative regression method. If the normality assumption is not met, these options will be considered, and the most appropriate will be implemented.

Additional preliminary multiple regressions will be done as needed to check the residuals once changes or transformations are made to one or more variables. Multivariate outliers will also be checked using Mahalanobis and Cook's distance. Cases that have a Mahalanobis distance score that exceeds the critical x^2 value at $\alpha = .001$ and/or have a Cook's distance score that exceeds 1.00 will be identified as potential multivariate outliers (Tabachnick & Fidell, 2018). If after further inspection, it is clear that these multivariate outliers are having a significant impact on the regression, these cases will be removed from the analysis.

Finally, the presence of multicollinearity will be evaluated using the variance inflation factors (VIF) values. Multicollinearity among independent variables can lead to large standard errors for the regression coefficients, causing them to not be significant. Cohen and colleagues

(2003) recommend the VIF values be no greater than 9. If this assumption is not met, independent variables will be excluded as needed.

Categorical variables will be dummy coded into mutually exclusive dichotomous variables with each variable coded 1 for "yes" and 0 for "no." Age, Gender, Social Class, and Parent Education Level are already dichotomous variables. They will each be recoded into new dummy variables with Traditional Student, Male, Not Low Income, and Continuing-Generation Student serving as the respective, excluded reference groups. For Ethnicity, five dummy variables will be created for Black/African American, Hispanic or Latino, Asian, American Indian, and Multiracial, and Caucasian will be used as the excluded reference group for each.

Methods

Sequential Multiple Linear Regression

A sequential multiple linear regression (MR) analysis will be used to answer the first research question. MR is used to determine if scores on a continuous dependent variable can be predicted based on scores on an independent variable. Using the variables included in this study, linear regression will be used to predict first-semester GPA (Dependent Variable) based on the measure of congruence (Independent Variable). This will show how a change in congruence is associated with a change in GPA. Furthermore, a sequential model will be used in order control for cultural considerations (Hinkle, Wiersma, & Jurs, 2003). Tabachnick and Fidell (2013) state that a sequential approach is appropriate when the goal is to hold identified variables constant in order to determine what a specific independent variable predicts "over and above" the initial set of variables (p. 133). In this case, gender, ethnicity, parent education level, and academic preparation will be included in the first block to control for their effects. The measure of congruence will then be included in the second block to see what it predicts over and above these cultural and precollege characteristics.

Sequential Logistic Regression

A sequential logistic regression (LR) analysis will be used to answer the second research question. The dependent variable in research question two is persistence into the second year, a dichotomous categorical variable. As Keith (2015) outlines, LR is an "appropriate analytic choice" when the dependent variable is a "naturally occurring categorical variable" (p. 227). Like the MR for the first research question, a sequential model is being used in order to determine if congruence predicts persistence into the second year over and above the identified cultural and precollege characteristics.

Limitations

There are several practical limitations to this study methodology that should be highlighted. As a single-site study, any findings will be limited to this specific institution. Applicability to other PTYCs will be limited. Furthermore, this is a cross-sectional study meaning that the data was collected at a single point in time. Thus, it is a snapshot and cannot account for how these variables and their relationships might change over time. The fact that the O*Net interest profiler was required before a student could have an advising appointment may have impacted the motivation and care that student had when completing it. Some student may have perceived the assessment as a box to be checked rather than an opportunity to learn about their occupational personality. Separate but related, there was also a lack of uniformity in the context in which students completed the instrument. While some may have completed it in a quiet environment, others may not have had this luxury. Students would often have to complete

PAWS in the waiting area prior to their appointment. Each of these final two points may have introduced measurement bias.

Summary

The purpose of this chapter was to provide an overview of the sources of data, dependent and independent variables, research methods, procedures of data cleaning, and the analytic techniques that will be used. Institutional data from a large, public two-year institution located in the southeast region of the United States will be used in this study. This includes precollege and cultural characteristics, IP Short Form results, first-semester courses, and academic outcomes for students in the sample. Precollege and cultural characteristics will be included as IVs in this study to determine if congruence, the IV that is the primary focus of this study, predicts academic outcomes over and above them. The IP Short Form results, first-semester courses, and DHOC will be used to calculate a single congruence score for each student. The DVs that will be used to represent academic success are first-semester GPA and persistence into the second year.

Two statistical methods will be used to answer two research questions. Sequential multiple linear regression will be used for the first research question. Precollege and cultural characteristics will be included in the first block and congruence in the second to predict first-semester GPA. For the second research question, sequential logistic regression will be used. The same IV's will be used and inserted in the same order to predict persistence into the second year. Prior to analysis, outliers will be identified and dealt with using a modification approach. Furthermore, the data will be assessed to determine if the linearity, multicollinearity, normality, homoscedasticity, and independence of errors assumptions are met. The results of these procedures and analyses are presented in Chapter Four.

CHAPTER 4

RESULTS

This dissertation explored the relationship between Holland's (1997) concept of congruence and student academic outcomes. In the context of this study, congruence refers to the fit between the student and their academic environment. In addition to improving the field's understanding of college departure at two-year college's, the methodology used in this study sought to address methodological issues with how previous studies had defined the "academic environment." Following Institutional Review Board (IRB) approval (see Appendix B), students' results on the O*Net Interest Profiler Short Form (IP Short Form) were obtained along with coursework, cultural and precollege characteristics, and academic outcomes from a large, public two-year college (PTYC) located in the southeast region of the United States.

Two research questions guided this study:

1. Is congruence between students' occupational personality type and their first-semester academic environment positively associated with first-semester GPA?
2. Does congruence between students' occupational personality type and their first-semester academic environment predict persistence into the second year?

Two statistical methods were used to address these questions with a sample of 1,257 PTYC students who had completed the IP Short Form. For the first question, a sequential, multiple linear regression (MR) analysis was used and a sequential logistic regression (LR) for the second.

The purpose of this chapter is to provide an overview of the results of these analyses. First, the preliminary analyses will be outlined. This includes the procedures that were used for data cleaning and assumption testing. The second section will detail the results of the primary analyses.

Preliminary Analyses

General Data Cleaning

Data was checked and cleaned in SPSS (version 24) using a combination of Tabachnick and Fidell's (2018), Keith's (2015), and Morrow and Skolits' (2017) guidelines. Frequencies were generated for each variable to check for coding issues, missing data, unexpected values, and any other initial concerns. Due to low group membership for Not Indicated (n = 1) American Indian (n = 2), Native Hawaiian or Pacific Islander (n = 1), and Asian (n = 19), the Ethnicity variable was recoded into a new variable in which these groups were combined into a single "Other" category (n = 27). As noted in the previous chapter, a number of values in the High School GPA variable were either beyond the 4.00 scale or missing entirely. These numbers ranged from 9.6 to 670, and after consulting with the institution's institutional research and records offices, it became clear that these values came from a combination of high school equivalency test (HiSET) scores and internal coding. With no way of converting these values to a 4.00 GPA scale, they were recoded as missing data. Values for remaining independent and dependent variables fell within expected ranges.

A Missing Variable Analysis (MVA) was conducted in SPSS to determine the scale and, if present, any patterns in this missing data. Univariate statistics showed that, of the 1,257 individuals included in the sample, 80 had missing high school GPA values or approximately 6.4% of the total sample. This is slightly larger than the 5% or less cutoff that Tabachnick and Fidell (2018) describe as "less serious" (p. 54). Further, separate variance *t* tests, showed that students with missing high school GPA data had higher first-semester GPAs than those with complete high school GPA data (p = .028), and the *p* value (.037) of Little's MCAR test was significant at the 0.05 level (Little & Rubin, 2002). Thus, the null hypothesis that the missing data are MCAR is rejected, and MNAR is inferred. For all these reasons, listwise deletion is not a viable option and mean estimation was used to impute missing values.

Due to a large number of missing values being the result of high school equivalency scores being available rather than traditional high school GPAs, it was theorized that a disproportionate number of these students would be classified as post-traditional students per the age variable. The high school GPA variable was recoded into a dichotomous, categorical variable with "Missing Data" and "Not Missing Data" serving as the two categories. Crosstabs was then used along with the Age variable to compare. Visual inspection revealed that Missing Data students were indeed disproportionately post-traditional students (73%) when compared to Not Missing Data students (13%). Based on this, a group high school GPA mean (2.71) was identified for available post-traditional students. The mean high school GPA for this subgroup was then imputed for students with missing high school GPA data. This high school GPA with imputed values was used moving forward in the study. **Table 4.1** shows the descriptives for the new high school GPA variable with imputed missing data.

Table 4.1
Descriptives for High School GPA with Imputed Missing Data

	N	Mean	Std. Deviation	Range
High School GPA	1257	3.02	0.58	0.79-4.00

Descriptives and standardized scores were used to determine if there were any univariate outliers among the continuous IV and DV variables. Outliers were defined as cases "with standardized scores in excess of 3.29" (Tabachnick & Fidell, 2018, p. 64). One negative outlier was identified for high school GPA. This value was modified to be one "unit" larger or smaller than 3.29 standard deviations from the mean. Per Tabachnick and Fidell (2018), this method allows for the scores to remain "deviant, but not as deviant as they were" (p. 67). One unit was defined as .1. No outliers were identified for congruence score or first-semester GPA.

Dummy Coding

Categorical data were dummy coded into mutually exclusive dichotomous variables with each variable coded 1 for "yes" and 0 for "no." Age, Gender, Pell Eligible, and Parent Education Level were already dichotomous variables. They were each recoded into new dummy variables with Traditional Student, Male, Not Low Income, and Continuing-Generation Student serving as the respective, excluded reference groups. For Ethnicity, dummy variables were created for Black, Hispanic, Two or more selected, and Other. White served as the excluded reference group.

Multiple Linear Regression Assumption Testing

A preliminary multiple regression analysis was conducted to create and inspect the scatterplot and histogram of the standardized residuals for the screening of multivariate normality, linearity, and homoscedasticity between predicted DV scores and errors of prediction. **Figure 4.1** shows the scatterplot of the standardized residuals for the preliminary analysis. In order for the assumptions of linearity and homoscedasticity to be met, the points should be randomly scattered around the horizontal axis with no identifiable pattern or shape. Visual inspection of **Figure 4.1** showed that this was not the case as the scatterplot of the residuals took

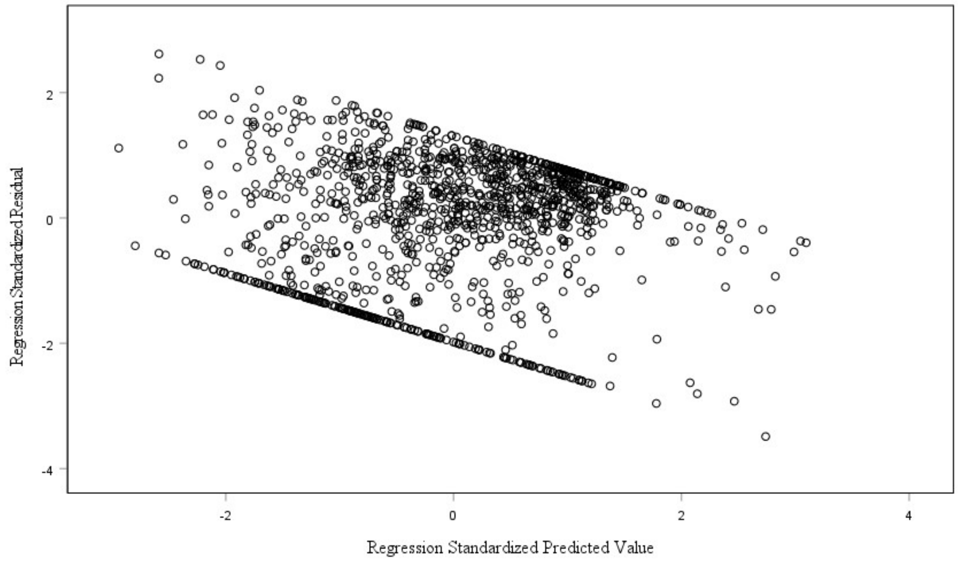

Figure 4.1
Scatterplot of the Standardized Residuals for First Preliminary Analysis

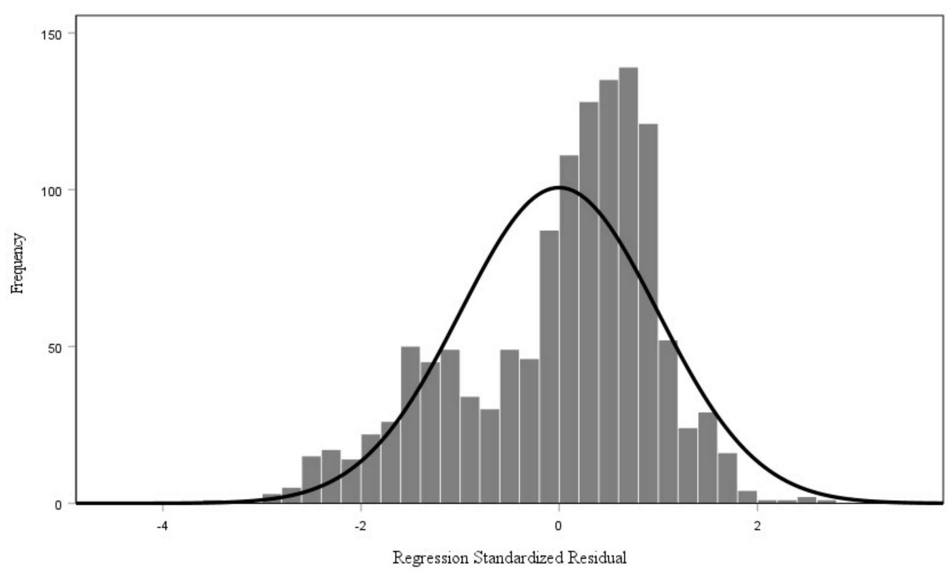

Figure 4.2
Histogram of Standardized Residuals for First Preliminary Analysis

on the shape of a tilted rectangle. Thus, multivariate homoscedasticity and linearity could not be inferred.

Furthermore, **Figure 4.2** shows the histogram of the standardized residuals for the preliminary analysis. Visual inspection indicated a significant departure from normality. More than just skewness and kurtosis issues, the distribution of the residuals appears to take on an almost bimodal shape. The violation of the multivariate linearity, homoscedasticity, and normality assumptions prompted a review of the univariate independent and dependent variables.

Histograms were generated for each continuous variable to check the shape of their distributions. **Figures 4.3 – 4.5** show the histograms for congruence score, high school GPA and first-semester GPA. Visual inspection of these figures showed that congruence score had a relatively normal distribution and that high school GPA had a negative skew. However, a quick glance at first-semester GPA, the dependent variable, immediately revealed a more pressing issue than skewness or kurtosis. Rather than resembling any kind of unimodal distribution, it had a u-shape due to the two values on each extreme end having a significant number of students, namely 0.00 and 4.00. Of the 1,257 members of the sample, 192 (15%) had a 0.00 GPA and another 134 (10%) had a 4.00 GPA. This distribution was undoubtedly a significant contributor to the issues observed when evaluating multivariate normality, linearity, and heteroscedasticity. Thus, a decision based on either statistical theory or literature in the field on what to do with this variable was necessary before proceeding any further in the analysis.

One option available was to convert the continuous first-semester GPA variable into a categorical variable and select an alternative statistical analysis, likely a multinomial logistic regression. While often treated as a continuous variable, an argument could also be made the GPA can be interpreted as a series of ordinal categories due to the fact that differences in GPA

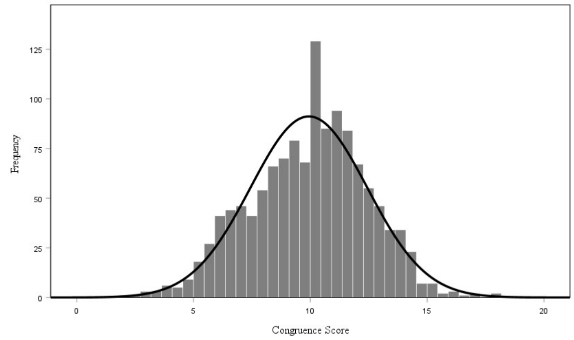

Figure 4.3
Histogram of Congruence Score

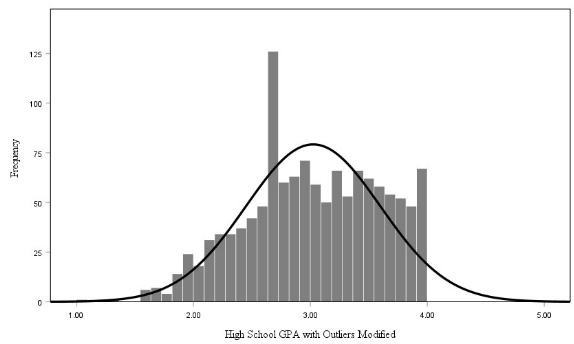

Figure 4.4
Histogram of High School GPA

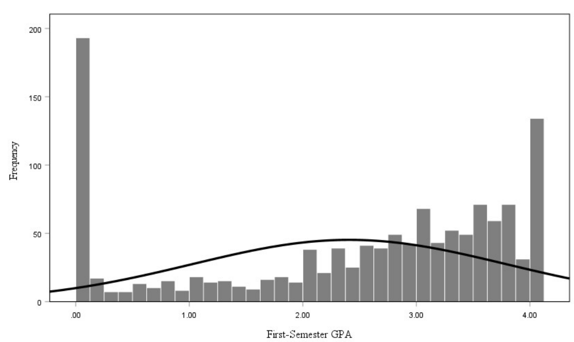

Figure 4.5
Histogram of First-Semester GPA

do not necessarily represent differences in performance (e.g., the difference between 1.00 and 1.50 may not be considered the same as the difference between 2.00 and 2.50). Tabachnick and Fidell (2018) do assert that logistic regression "is more appropriate than multiple regression when the distribution of responses over a set of categories seriously departs from normality, making it difficult to justify using an ordered categorical variable as if it were continuous" (p. 369).

However, considering GPA a series of ordinal categories would be to assert that every point on a two-decimal, 4.00 GPA scale is a category. This would be 401 categories. Of course, you would not run a multinomial logistic regression with this many categories in the independent variable. Rather, you would group these into a much smaller range of categories (e.g., 0.00-1.00, 1.01-2.00, 2.01-3.00, and 3.01-4.00) due to the precision of the GPA variable. It is this precision that leads it to be treated as a continuous variable as many studies do, and when it comes to categorizing a continuous variable, Keith (2015) states that "it is generally not a good idea to turn a continuous variable into a categorical one" (p. 226). He elaborates that categorizing a continuous variable "reduces-throws away! wastes!-its variance and reduces its correlation with any other variable" (p. 226). He concludes this point by agreeing with Thompson (2006), who called this "data mutilation" (as cited in Keith, 2016, p. 226). For this reason, the decision was made to continue with a multiple linear regression and explore other methods for resolving this issue.

Keith (2015) recommends either transformation of the dependent variable or eliminating subgroups as potential strategies for dealing with excessive nonnormality of residuals. Log transformation is recommended for distributions that diverge significantly from normality (Tabachnick & Fidell, 2018). The Compute Variable function was used in SPSS to conduct a log transformation of the dependent variable, first-semester GPA. Tabachnick and Fidell (2018) emphasize that "it is important to check that the variable is normally or nearly normally distributed after transformation" (p. 75). **Figure 4.6** shows the histogram for the transformed dependent variable.

This histogram reveals that the log transformation did not fix the issues of normality. This makes sense as data transformations aim to fix issues of skewness. The issue with the original distribution is its multimodality at both ends. Thus, transformation was not a viable solution. Keith's (2015) recommendation to eliminate subgroups was explored, requiring insights from the literature and a deeper understanding of subgroup impacts on the distributions.

One of Holland's (1997) basic assumptions is that the outcomes of congruence are "determined by the interaction between personality and environment" (p. 4). The key word

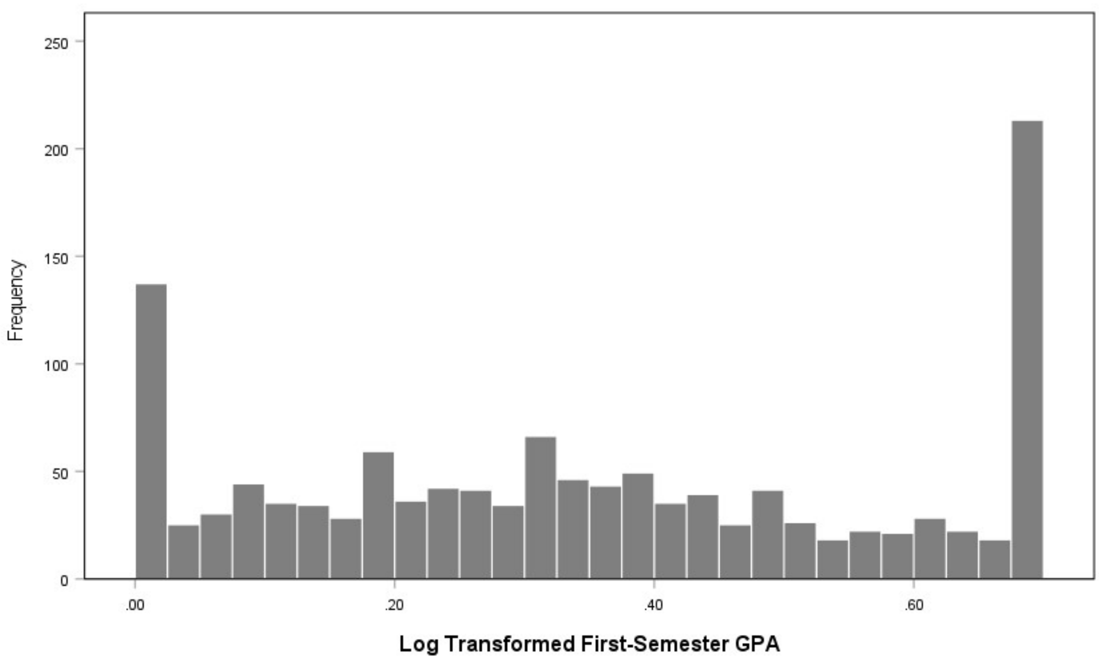

Figure 4.6
Histogram of Log Transformed First-Semester GPA

in that statement is *interaction.* An interaction assumes that the individual is present and participating in the environment. A semester GPA of 0.00 suggests that the student has likely disengaged either entirely from their classroom environment or at least to the extent that no reciprocal interaction has truly occurred. This is not to say that incongruence at the classroom level may not play a part in that disengagement, but continuing to include these students could undermine a central tenet of the theoretical framework and mask the significance of congruence as a predictor variable. Thus, the decision was made to exclude students who earned a 0.00 semester GPA, reducing the sample size by 192 students to 1,065.

A second preliminary analysis was done to assess the normality of the standardized residuals. **Figures 4.7 – 4.8** show the updated histogram and a P-P plot with students who earned 0.00 first-semester GPA removed from the sample. While the removal of this subgroup did help eliminate the bimodal shape of the distribution, the histogram appeared to be leptokurtic and negatively skewed. Reviewing the P-P plot confirmed that the normality assumption had not yet been satisfied to the necessary extent to continue with the analysis. Furthermore, first-semester GPA was still the likely culprit. The bunching up values at the extreme right end of the distribution, 4.00, still created a significant negative skew.

A log transformation of first-semester GPA with students who earned 0.00 removed was done and included in a third preliminary analysis. Note that the log transformed variable was re-reflected for ease of interpretation. While **Figures 4.9** and **4.10** show that this dramatically improved the shape of the histogram and P-P plot of the standardized residuals, **Figure 4.11** continued to show that there was a heteroscedasticity issue. Furthermore, even when transformed and re-reflected, first-semester GPA with students who earned 0.00 removed still had an

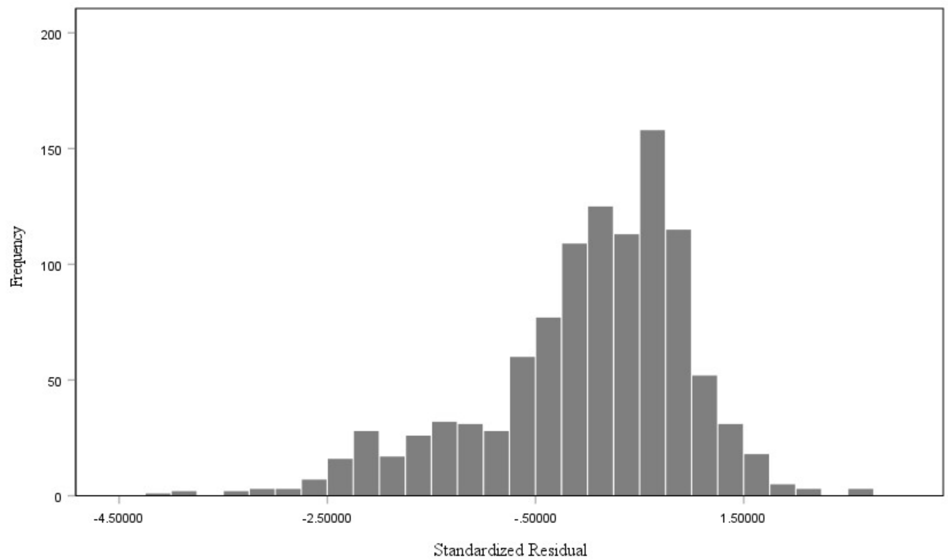

Figure 4.7
Histogram of Standardized Residuals for Second Preliminary Analysis

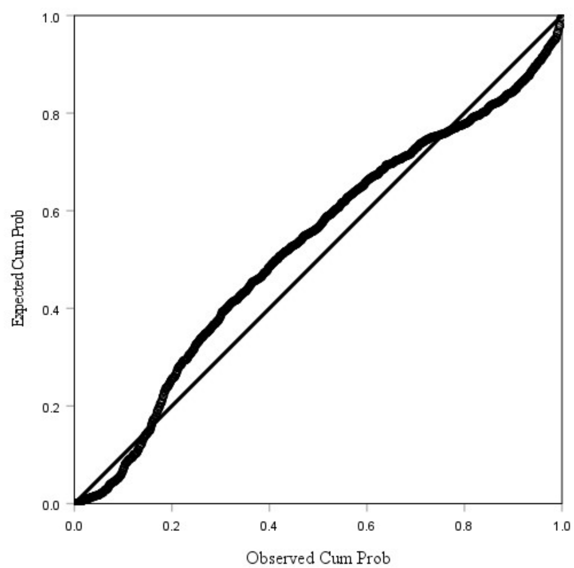

Figure 4.8
P-P Plot of Standardized Residuals for Second Preliminary Analysis

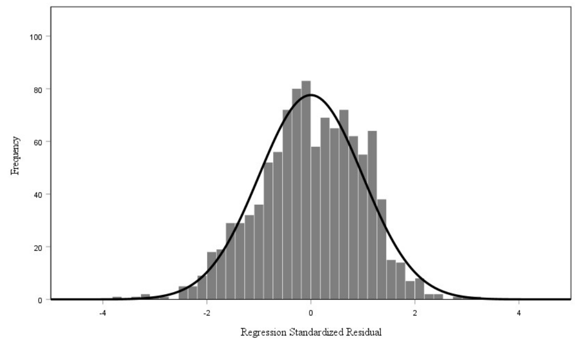

Figure 4.9
Histogram of Standardized Residuals for Third Preliminary Analysis

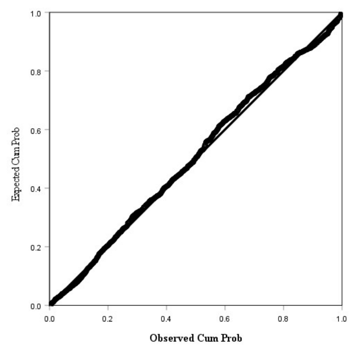

Figure 4.10
P-P Plot of Standardized Residuals for Third Preliminary Analysis

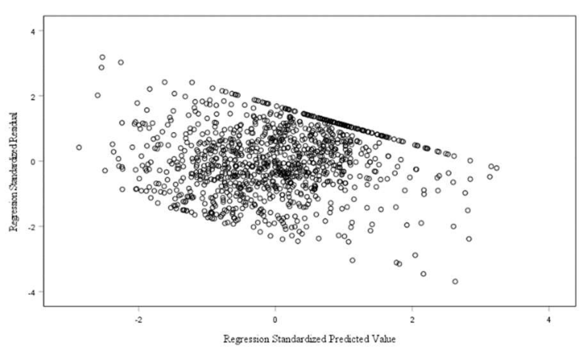

Figure 4.11
Scatterplot of the Standardized Residuals for Third Preliminary Analysis

alarming histogram shape as seen in **Figure 4.12**. Log transformation had not corrected the univariate shape of the distribution, and this was certainly due to the large number of 4.00 GPAs.

At this point, the decision was made to once again consult Holland (1997) to find a path forward. As highlighted in previous chapters, Holland asserted that his theory "cannot be applied successfully without the observance of a few boundary conditions" (p. 13). Chief among them is that the notion of "other things being equal…" should be at the forefront of any methodology utilizing his theory. While this study sought to control for prominent precollege characteristics, the first-semester outcomes of this sample demonstrate just how much of an "ill-structured problem" colleges student performance and departure is (Braxton et al.,2004) as there is no subgroup or combination of subgroups (e.g., part-time, post-traditional, first-generation, etc.) that could be removed that would make this variable distributed normally. The reality is that there are a large number of intersecting reasons for why some students are successful and others are not.

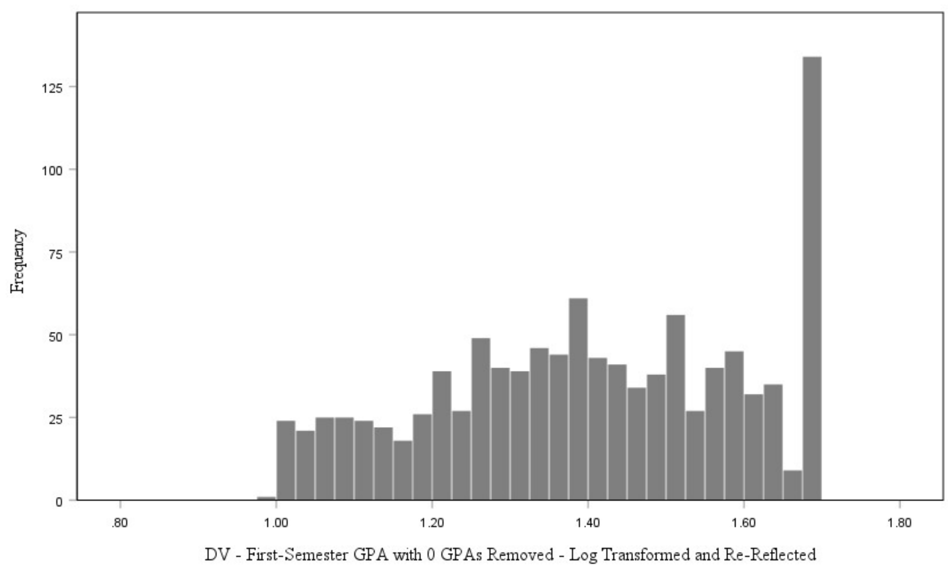

Figure 4.12
Histogram of Transformed and Re-Reflected First-Semester GPA with 0.00 Removed

In the absence of this simplicity, the choice was made to extend the logic that was used to remove students who earned a 0.00 GPA. If a 0.00 GPA across all classroom environments represents complete disengagement due to one or more factors above and beyond their congruence with those environments, then a student with a 4.00 GPA could be inferred to have had one or more factors that *maximized* their ability to earn all A's in their coursework above and beyond their congruence.

While it is entirely possible that congruence is among these factors, the inclusion of these students could obscure its potential predictive power and any insights that these analyses might provide. Thus, the decision was made to also remove students what had earned a 4.00 GPA in their first semester as well. The obvious criticism is that their exclusion may inflate the predictive power of the included independent variables. This criticism is valid and will be acknowledged as a limitation of the study and an area for further exploration should congruence be found to have a significant relationship with first-semester GPA.

Removing students with 4.00 GPAs further reduced the sample size to 930 students. The updated histogram for first-semester GPA can be found in **Figure 4.13.** Because it was still negatively skewed, a log transformation was done using the Compute Variable function and the formula *LG10 (K - X)* where *K* is a constant that is equal to the largest score plus one and X is the actual value. Because the variable is negatively skewed, the actual value must be subtracted from the constant in order to reflect the variables before a transformation was done. For ease of interpretation, the transformed variable was then re-reflected for use in the analyses.

This transformed first-semester variable was used in a fourth preliminary regression to generate a new scatterplot, P-P plot, and histogram for the standardized residuals as seen in **Figures 4.14 – 4.16**. Visual inspection of these updated graphs showed that normality

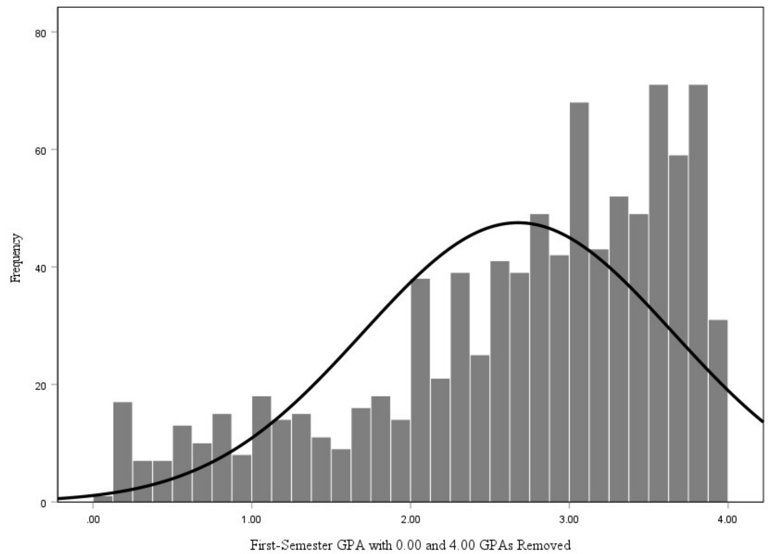

Figure 4.13
Histogram of First-Semester GPA with 0 and 4.00 GPAs Removed

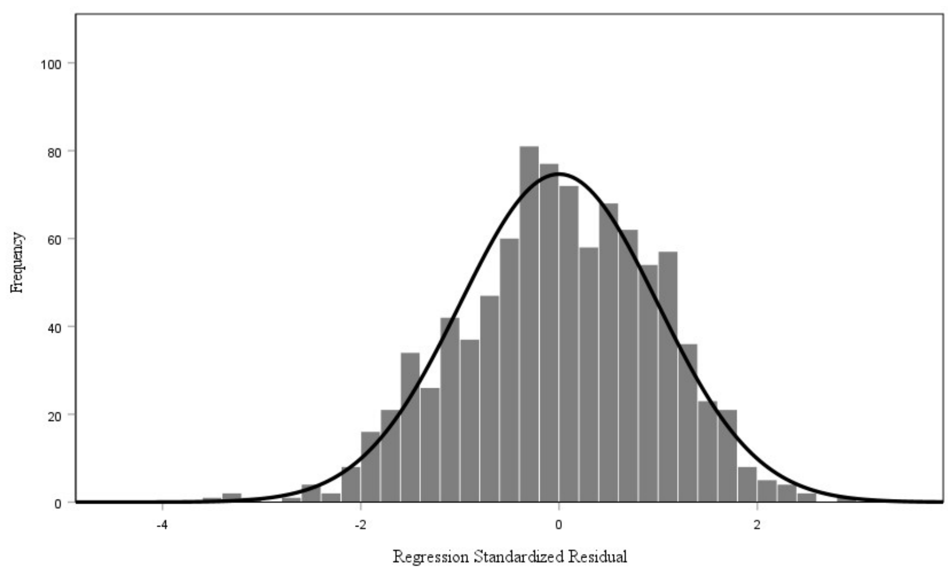

Figure 4.14
Histogram of the Standardized Residuals for the Fourth Preliminary Analysis

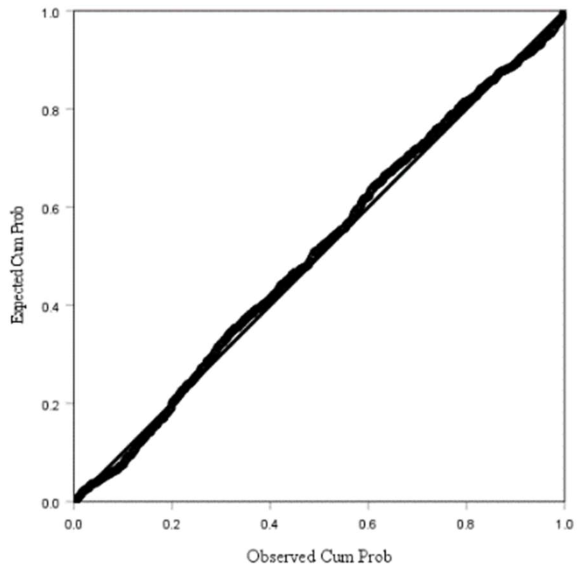

Figure 4.15
P-P Plot of Standardized Residuals for the Fourth Preliminary Analysis

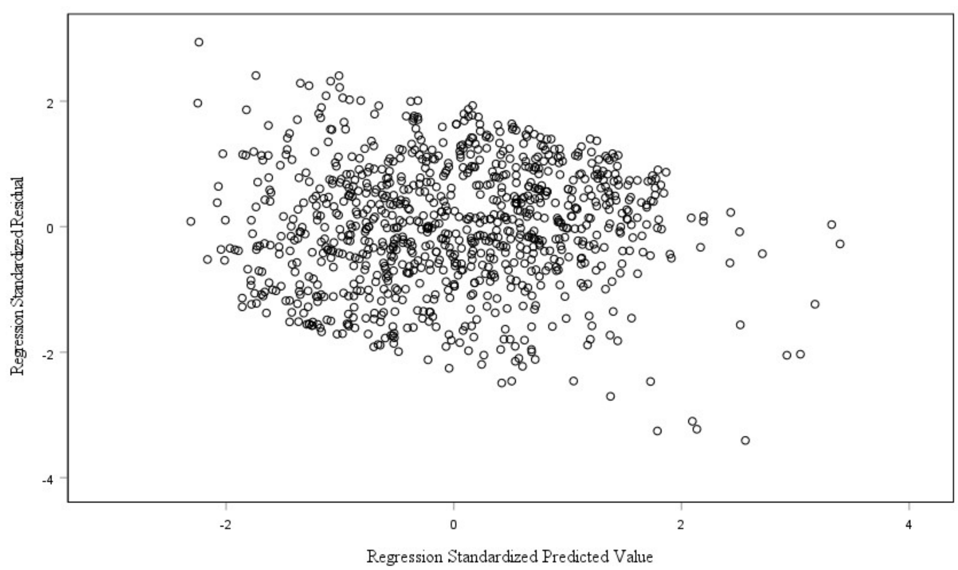

Figure 4.16
Scatterplot of the Standardized Residuals for the Fourth Preliminary Analysis

assumption continued to be met based on the histogram and P-P plot. While the scatterplot appears to have improved significantly, it had still somewhat maintained the tilted rectangle shape observed in the initial preliminary regression, pointing to a potential violation of the homoscedasticity assumption. Tabachnick and Fidell (2018) note that heteroscedasticity "does not invalidate the analysis so much as weaken it" (p. 108).

This shape is largely attributable to a small number of discrepant values on both ends. Thus, an evaluation of multivariate outliers was conducted both as a standard part of the data cleaning process and to inform the study on how impactful these values might be. Discrepancy measures the extent to which a case is in line with the others (Tabachnick and Fidell, 2018). In other words, it is how far the point is from the Y axis. Keith (2015) recommends evaluating datapoints that have a residual greater than the absolute value of 2. Thirty datapoints fell above or below 2 on the Y-axis and four of these had an absolute value that exceeded 3. Mahalanobis and Cook's distance scores were generated for these 30 datapoints to further assess them. The critical x^2 at $\alpha = .001$ for 10 degrees of freedom is 23.209. Reviewing the Mahalanobis scores showed that several of these datapoints did exceed this threshold. However, no Cook's score came close to exceeding 1, suggesting that the actual influence of these potential multivariate outliers was minimal. For this reason, these data points were retained in the sample. Collinearity statistics showed a variance inflation factors (VIF) value of 1.001, indicating that multicollinearity was not an issue in the analysis.

Logistic Regression Assumption Testing

As noted by Mayers (2013), "there are fewer assumptions and restrictions for logistic regression compared with multiple linear regression" (p. 447). The absence of multicollinearity among the predictor variables is necessary, but as was just established in the previous section, this is not an issue among this set of predictors. What is *not required* is that the variables be normally distributed. Thus, there is no need to transform non-normally distributed predictor variables or exclude subgroups due to non-normality in the dependent variable. For these reasons, all 1,257 cases were included in the logistic regression to answer the second research question.

Primary Analyses

Research Question 1

Does congruence between students' occupational personality type and their first-semester academic environment predict first-semester GPA?

To address the first research question, a sequential multiple linear regression was conducted using first-semester GPA as the dependent variable (DV) and congruence as the (IV) to assess how a change in congruence is associated with a change in GPA. To strengthen this analysis, a sequential model was used control for cultural and precollege characteristics. Gender, age, ethnicity, Pell eligible, first-generation status, and high school GPA were included in the first block to control for their effects. The measure of congruence was then included in the second block to see what it predicts over and above these cultural and precollege characteristics.

After step 1, the model was found to be significant, $F(9, 930) = 24.716$, $p < .001$ and accounted for 19.5% of the variance in first-semester GPA per the R-Squared value. Among the variables, high school GPA ($\beta = .386$, $p < .001$), being a post-traditional student ($\beta = .274$ $p < .001$), and being a female ($\beta = .077$, $p = .012$) were significant predictors at a significance level of .05. The various ethnicity variables, first-generation status, and being Pell Grant eligible did not significantly contribute to the model. **Table 4.2** shows the overall results for the model and each variable.

Table 4.2
Model 1 Regression Results

Predictors	β	t	Sig.
High School GPA	0.386	11.941	<.001
Post-Traditional	0.274	8.777	<.001
Female	0.077	2.507	0.012
Pell Grant Eligible	0.028	0.904	0.366
First-Generation	-0.037	-1.179	0.239
Other Race	0.005	0.162	0.871
Black	-0.038	-1.23	0.219
Hispanic	-0.009	-0.307	0.759
Two or More Races	-0.045	-1.512	0.131

$F(9, 930) = 24.716$, $p < .001$, $R = 0.441$, $R^2 = 0.195$, Adj. $R^2 = 0.187$

Congruence Score was added in step 2. **Table 4.3** shows the overall results for the model and each variable. The second model remained significant, $F(10, 930) = 22.276$, $p < .001$, with no change in accounted variance, 19.5%, in First-Semester GPA. High school GPA ($\beta = .386$, $p < .001$), being a post-traditional student ($\beta = .274$, $p < .001$), and being female ($\beta = .073$, $p = .022$) remained significant predictors. Traditionally, you would interpret coefficients in a multiple regression as "how outcome values change for every unit change in the predictor" (Mayers, 2013, p. 402). However, when a dependent variable has been log-transformed, it is more appropriate to interpret the coefficient as a percentage change using the formula, *Exp*(β) – 1 where "exp" represents exponentiate (Monk, 2000). In the case of high school GPA, the formula would be:

Exp(.386) – 1

1.471 – 1

.471

Table 4.3
Model 2 Regression Results

Predictors	β	t	Sig.
High School GPA	0.386	11.950	<.001
Post-Traditional	0.274	8.780	<.001
Female	0.073	2.298	0.022
Pell Grant Eligible	0.028	0.886	0.376
First-Generation	-0.036	-1.149	0.251
Other Race	0.005	0.175	0.861
Black	-0.038	-1.248	0.212
Hispanic	-0.008	-0.262	0.793
Two or More Races	-0.045	-1.514	0.130
Congruence Score	0.020	0.669	0.504
$F(9, 930) = 22.276$, $p < .001$, $R = 0.442$, $R^2 = 0.195$, Adj. $R^2 = 0.186$			

Converted to a percentage, that would be 47.1%. Thus, for every one-unit increase in high school GPA, first-semester GPA increased by about 47.1%. Furthermore, Congruence Score ($\beta = .018$, $p = .556$) did not contribute significantly to the regression, indicating that this variable does not provide additional predictive value for First-Semester GPA.

Research Question 2

Does congruence between students' occupational personality type and their first-semester academic environment predict persistence into the second year?

For the second research question, a sequential, binary logistic regression was used to assess if congruence (DV) is a significant predictor of student persistence into the second year (IV). Like the analysis used for the first research question, a sequential approach was used to see if congruence improved prediction beyond gender, age, ethnicity, parent education level, and high school GPA. Tabachnick and Fidell (2018) state that a simple option for sequential logistic regression is to do "multiple runs, one for each step of the proposed sequence" and then evaluate the difference between the models (p. 360). For this research question, two runs were conducted, Model 1 and Model 2. Model 1 contained the cultural and precollege characteristics, and Model 2 added congruence score.

Model 1 was statistically significant, $x^2(9, N = 1257) = 110.26$, $p < .001$, indicating that this initial set of predictors more effectively distinguished which students persisted into the second year when compared to the baseline. This model explained between 8.4% (Cox & Snell R Square) and 11.2% (Nagelkerke R Square) of the variance in the dependent variable and correctly classified 63.2% of cases. Model 2 added congruence score as a predictor variable. **Table 4.4** displays the overall evaluation of Model 2, and **Table 4.5** shows the regression coefficients, Wald statistics, degrees of freedom, odds ratios, and 95% confidence intervals for odds ratios for each of the predictors. While it remained statistically significant, $x^2(10, N = 1257) = 110.89$, $p < .001$, the addition of the congruence score variable did not increase its predictive power. The explained variance remained approximately the same, and the percentage of correctly classified cases actually decreased slightly to 62.9%.

The Hosmer and Lemeshow test (HLT) is a goodness-of-fit test that compares the observed and expected frequencies of the outcome variable to assess how well the model fits the observed data. A significant HLT statistic indicates that there is a difference between the observed and expected frequencies, indicating that the model does not fit the data well. The HLT statistic for this statistic ($x^2 = 8.065$, $p = .427$) was not significant, suggesting that the final model fit the model well.

Per the Wald criterion, high school GPA, $x^2(1, N = 1257) = 81.7$, $p < .001$, being a post-traditional student, $x^2(1, N = 1257) = 12.3$, $p < .001$, and, by a very small margin, being Black, $x^2(1, N = 1257) = 4.35$, $p = .04$, were the only significant contributors to the model.

Table 4.4
Overall Evaluation of the Model 2 Logistic Regression

	x^2	df	Sig.	Cox & Snell R^2	Nagelkerke R^2	Overall Correct %
Omnibus Test	110.89	10	<.001	0.084	0.113	62.9
Hosmer & Lemeshow	9.50	8	.301			

Table 4.5
Logistic Regression Predicting the Likelihood of Persisting into the Second Year

	B	S.E.	Wald	df	Sig.	OR	95% C.I.	
							Lower	Upper
Female	-0.16	0.13	1.60	1	0.21	0.85	0.66	1.09
Other Race	-0.34	0.40	0.69	1	0.41	0.72	0.33	1.57
Black	-0.53	0.25	4.35	1	0.04	0.59	0.36	0.97
Hispanic	0.23	0.23	1.03	1	0.31	1.26	0.81	1.97
Two or More Races	-0.33	0.32	1.05	1	0.31	0.72	0.38	1.35
First Generation	-0.12	0.13	0.89	1	0.35	0.89	0.69	1.14
High School GPA	1.07	0.12	81.70	1	<.001	2.91	2.31	3.67
Pell Grant Eligible	0.21	0.13	2.71	1	0.10	1.24	0.96	1.59
Post-Traditional	0.59	0.17	12.30	1	<.001	1.81	1.30	2.51
Congruence Score	0.02	0.02	0.63	1	0.43	1.02	0.97	1.07
Constant	-3.22	0.442	53.282	1	<.001	0.04		

Summary

This chapter provided an overview of the results of the methodological procedures detailed in Chapter Three. Sequential multiple linear regression and sequential logical regression were used to answer two research questions about the relationship between congruence and academic outcomes. The chapter began with a description of the preliminary analyses conducted to clean the data and test assumptions. Data cleaning involved checking for coding issues, missing data, and unexpected values. Missing data analysis revealed that 6.4% of the sample had missing high school GPA values, which required imputation using mean estimation due to the presence of non-random missingness. Outliers were identified and modified for the high school GPA variable. Dummy coding was performed for categorical variables, and data transformation was applied to address violations of normality for high school GPA and first-semester GPA.

A preliminary regression was run to review the assumptions of multivariate normality, linearity, and homoscedasticity through scatterplots and histograms of standardized residuals. Further examination revealed issues with the distributions of high school GPA and first-semester GPA, with the latter showing a bimodal distribution due to a substantial number of students with GPAs of 0.00 and 4.00. Considering the complex nature of student departure at two-year colleges and the absence of congruence among the factors previously identified in the literature, it was decided to exclude students with extreme GPAs from the analyses to avoid confounding effects. Logarithmic transformations were used for both this new first-semester GPA variable and high school GPA.

For the first research question, a sequential multiple linear regression was used to determine if congruence score significantly predicted first-semester GPA over and above several precollege and cultural characteristics. Similarly, a sequential logistic regression was used for the second research questions to see if congruence score did the same for persistence into the second academic year. While the final models for both the sequential multiple linear regression and logistic regression were significant, congruence score was not a significant predictor for either analysis.

CHAPTER 5

DISCUSSION AND CONCLUSION

This study was conducted with the goal of contributing to the literature on student performance and persistence. While there has been a great deal of research and theory development when it comes to this topic, there were some clear gaps that warranted attention. To begin with, public two-year colleges (PTYCs) do not receive enough focus from the field despite the significant role they play in the United States' higher education system and the concerning student outcomes that they produce. Furthermore, a disproportionate amount of emphasis has been placed on student characteristics rather than the academic environment. What is needed, then, is the application of a theoretical framework that equally considers both the individual and their environment when predicting outcomes. Personality-environment fit (P-E fit) theories meet this criteria, and John Holland's (1997) theory of vocational personalities and work environments is unarguably one of the most widely researched P-E fit theories available. Holland posited that congruence between an individual and their environment leads to increased satisfaction, performance, and stability. In the college setting, performance is best measured by GPA and stability by persistence. Thus, this study was guided by two research questions:

1. Does congruence between students' occupational personality type and their first-semester academic environment predict first-semester GPA?
2. Does congruence between students' occupational personality type and their first-semester academic environment predict persistence into the second year?

To answer these questions, this quantitative study utilized a final sample of 1,257 students who had completed a version of the O*Net Interest Profiler Short Form (IP Short Form). The results included the rank order and scores for all six RIASEC types for each student.

Summary of Results

Research Question 1

For the first research question, a sequential multiple linear regression was conducted using first-semester GPA as the dependent variable (DV) and congruence as the (IV) to assess how a change in congruence is associated with a change in GPA. Age, gender, ethnicity, Pell eligible, first-generation status, and high school GPA were included in the first step as controls. The measure of congruence was then included in the second step to see what it predicts over and above these cultural and precollege characteristics.

While the overall model was significant after step 1, only three variables were significant predictors of first-semester GPA. Two of these variables, high school GPA and gender were consistent with previously cited literature. The results showed that students with a higher high school GPA tended to have higher first-semester GPAs, and that women, on average, have a higher first-semester GPA when compared to men. Better academically prepared students have a greater chance of being successful in college and women graduate at higher rates compared to men (Cohen, Brawer, & Kisker, 2013; Skomsvold, 2014). However, the third variable, age, was in contrast with what was expected.

On average, post-traditional students had higher first-semester GPAs when compared to traditional students. Past research has shown that traditional students have higher associate degree completion rates when compared to post-traditional students (Hoachlander et al., 2003). Granted, first-semester GPA and graduation are different outcomes. Perhaps the challenges that cause post-traditional student dropout occur over the course of several semesters and are less related to academic ability and more to the demands of being a post-traditional student. Another consideration is the role that TN Reconnect might have played in changing the post-traditional student demographics and outcomes in Tennessee PTYCs. As discussed in Chapter One, TN Reconnect is a last-dollar scholarship that covers the tuition for any post-traditional student, defined as 24 or older, who does not have a postsecondary degree and enrolls in a Tennessee community college. Fall 2018, the semester under inspection, was the first semester in which this scholarship was available for TN residents. In addition to attracting more post-traditional students to attend a PTYC, the financial relief that comes from having tuition covered may have led to a meaningful increase in post-traditional student success.

Congruence score, defined as the congruence between the individual student and their first-semester coursework, was entered in the second step. It did not significantly add to the model. Given the novelty of classifying the academic environment as the actual classes that the student is taking rather than their academic major, it is difficult to compare this to previous

literature. The closest analogue is interest-major congruence. In this respect, the body of literature is fairly mixed with some researchers finding that interest-major congruence does adds to the prediction of GPA, including first-semester, second-year, third-year, and at graduation (Rocconi et al., 2020; Tracy, Allen, & Robbins, 2012; Tracy & Robbins, 2006). That being said, Allen and Robbins (2010) found that interest-major congruence *was not* a significant predictor of first-year academic performance. This article is notable as it is the only identifiable study that included a two-year sample.

Research Question 2

To answer the section research question, a sequential, binary logistic regression was used to assess if congruence (DV) is a significant predictor of student persistence into the second year (IV). Like the analysis used for the first research question, the first step included age, gender, ethnicity, Pell eligible, first-generation status, and high school GPA control for their effects. Once again, the model after step 1 was significant, but high school GPA was the only significant predictor. This was a little surprising given the literature on unequal outcomes among these groups. Both women and traditional students have much higher degree completion rates when compares to men and post-traditional students respectively (Cohen et al., 2013; Hoachlander, Sikora, & Horn, 2003) Likewise, Black students have a much lower retention rate when compared to other ethnicities (NSC Research Center, 2018).

Congruence score was added in the second step and did not meaningfully contribute to the model. As was concluded for the first research question, the congruence between a student and their first-semester classes does not appear to be a significant predictor of this academic outcome. The literature on the relationship between interest-major persistence is less mixed. Tracy and Robbin (2006) found that it was not a significant predictor of first- and second year enrollment or graduation, and Tracy, Allen, and Robbins (2012) likewise concluded that it did not add to the prediction of second- and third-year persistence. Allen and Robbins (2010) did find that interest-major congruence to have a significant relationship with degree attainment for four-year institutions but *not* two-year. It appears that interest-class congruence, an even more precise measure of congruence that allows for the inclusion of undecided students, is in line with the broader range of outcomes that other researchers have found.

Several hypotheses and limitations for why interest-classroom congruence was not significant in either analysis will be discussed in the next section. However, this initial study indicates that congruence between a student and their first-semester classes is not a notable factor for neither predicting first-semester GPA nor second year persistence

Discussion

From conception to implementation, this study was guided by a simple idea. The more a student has fit with their first-semester academic environment, the more successful they will be. As this idea was developed, it became clear that the road to testing this idea was not clear of methodological and theoretical issues. The next few sections of this chapter will highlight and summarize some of the more prominent concerns and provide recommendations for future research for each.

General Limitations and Recommendations

Limitations. As anticipated and outlined in Chapter Three, there were several general, practical limitations to this study. First, as a single-site study, the results are specific to the public two-year college (PTYC) where the research was conducted. Generalizability to other PTYCs is limited due to differences in student demographics and institutional structure. Second, data were collected at one moment in time, making it a cross-sectional study that does not allow for any analyses of these variables and how relationship might change over time. Third, completion of the instrument was mandatory before students could attend their advising appointment, which may have introduced measurement bias due to some students completing it as a formality. Finally, students completed the assessment in a variety of conditions. Many students would complete the assessment on laptops in the Advising waiting area before their appointment. This may have also introduced measurement bias.

Recommendations. These practical limitations lead to practical recommendations for future studies. Multi-site and longitudinal research designs might reveal patterns in the variables and relationships that this study may have missed. While a much larger undertaking, combining UNIACT data with multi-institutional data on first-semester courses, first-semester GPA, and fall-to-fall retention over the course of several years is one potential design that would allow for much greater generalizability. This could also address the measurement bias introduced by the lack of a uniform context as students are required to take the ACT in a controlled environment. While the ACT is still a procedural formality for many students, it could be argued that its direct relevance to their goals is much clearer in their minds, resulting in a more genuine completion of its requirements and reduction in potential measurement bias.

Classifying the Academic Environment Revisited

Limitations. One possible methodological issue leading to the lack of significance may be found in the ongoing difficulty of classifying the "academic environment." A foundational assertion that motivated this study was that a college student's chosen program of study is both an insufficient variable and an unreliable representation of their academic environment in the first semester. It is insufficient in that it forces the higher education researcher to exclude undecided students and unreliable due to the heterogeneous environment that college students, especially first-semester students, experience. While the academic environment becomes more homogeneous as a student completes their general education requirements and focuses on major-specific coursework in the latter years of their college experience, these students have already demonstrated performance and stability in the environment and there is no opportunity to study those who have dropped out. These issues are even more pronounced for community college students as almost 40% do not return for the second year (McFarland et al., 2018).

This study attempted to address this issue by going a step further and categorizing their sub environments (i.e., classes) based on the larger instructional program. However, this may still not be sufficient for multiple reasons. As described in Chapter Two, occupational environments are typically classified in one of three ways: (1) the incumbent method, (2) the empirical method, and (3) the judgment method. The incumbent method classifies environments based on the individuals who make up that environment, the empirical method uses occupational analysis data and classification equations to assign RIASEC types to an environment, and, finally, the judgment method requires a manual review by a trained judge to classify an environment.

Unfortunately, the classification of majors has not historically received the same level of attention as occupations. The *Dictionary of Holland Occupational Codes (DHOC)* includes the only publicly available, comprehensive classification of instructional programs using the RIASEC types (Gottfredson & Holland, 1996). Not only is this resource nearly thirty years old, but it also uses none of the aforementioned methods to classify majors. Rather, instructional programs are assigned RIASEC types based on what could be called the association method. Here, a crosswalk is used between a dataset of majors and a dataset of careers that are associated with that major to categorize each academic program based on the existing classifications of those careers. This strategy has some potential methodological issues.

First, many majors do not have a straightforward relationship with a specific career or group of careers. For example, take a major like Philosophy which the DHOC (1996) classified as AIE. While this author could not locate a publicly available dataset that shows which occupations were used to make this classification, O*Net provides a more recent crosswalk between the 2018 Standard Occupational Classification system (SOC), including each occupation's RIASEC profiles, and the 2020 Classification of Instructional Programs (CIP). Thus, it is possible to inspect the careers that are currently associated with a specific major.

Interestingly, the SOC-CIP crosswalk shows that both "Philosophy" and "Philosophy and Religion" instructional programs are only associated with the occupation "Philosophy and Religion Teachers, Postsecondary" which has a Holland Code of SAI. However, the U.S. Bureau of Labor and Statistics (2022) reports a variety of top-employing occupational titles for individuals with a Philosophy and Religion degree that include lawyer (EIA), clergy (SEA), and software developer (ICR for "Computer Programmer"). Not only are these occupations not connected to the instructional program per the crosswalk, but they are very different professions. One would not typically think of clergy and software developer as aligning with the same personality type, and indeed, their Holland codes validate this intuition as they represent two different halves of the Holland Hexagon. This all points to the conclusion that the association method of classifying instructional programs is, at best, outdated in its methodology and, at worst, not valid at all. Considering how widely the DHOC has been used over the last 18 years to help students make choices about their studies, this is very concerning.

Some more contemporary studies have avoided this issue by developing environmental profile scores that can be then used in the calculation of a congruence score (Allen & Robbins, 2010; Tracy, Allen, & Robbins, 2012; Tracy & Robbins, 2006). However, this application of the incumbent method requires access to a large data set of individuals who have demonstrated success in that instructional program. While this is theoretically possible at four-year institutions, researchers would have to have access to the complete RIASEC of graduated students from every program to then develop profiles that can be compared to incoming students.

It is not realistic at two-year institutions where the enrollment patterns are far more sporadic and instructional programs far less clear.

Furthermore, this study not only relied on the problematic classification of majors based on their association with careers but also on the assumption that specific classes could be classified based on their association with the larger instructional program to which they belong. However, even this is questionable. Take a course like MUS 1030, Music Appreciation. This common general education course was taken by 187 of the students in the sample or almost 15%. Because it is associated with the instructional program "Music, General," it was assigned the Holland Code AES. At first glance, this seems appropriate. After all, music is one of the arts and allows for the creativity that a person with Artistic as a dominant type would want. However, an inspection of the Master Syllabus for this course provides the following catalog description:

> *Developing listening skills and an understanding of Western music from the ancient world through the 20th century.*

Elsewhere in the syllabus, the course is self-described as a "music history-emphasis course." Indeed, a review of the topic for each week shows that after a brief overview of the "elements of music," each subsequent week focuses on a historical period. Perhaps the most revealing of all is the section on how students are evaluated. 85% of the final grade is determined through "Testing Procedures" that may "include multiple choice, matching, short-answer and/or discussion questions."

The point being made is that, despite being a "music" class, nowhere in this syllabus is there mention of assignments or activities that allow for creativity, self-expression, or the actual creation of art. Ultimately, it is what it describes itself as, a history class. Notably, Artistic is not even in the top three for the instructional program classification of history which has an assigned

Holland code of ESR. This is not to say that an Artistic individual who loves music would not have a greater interest in this specific area of history, increasing the chances of success. However, it still stands that this course is not an environment that requires and rewards the most distinctive characteristics of the Artistic individual. This point could be made for a variety of general education classes that are taken at the beginning of a student's college career. Nearly all Humanities electives are "History of" classes, and the Social Sciences largely focus on theoretical helping relationships rather than participating in actual ones. Of course, this begins to get into the very nature of the academic environment and the tension between theoretical and applied learning. It is not the goal for this discussion to tread too deeply into those waters, but only to acknowledge that, despite this study's attempt to classify and measure the academic environment, there continue to be unsolved issues that may be masking the role that congruence plays in student retention and performance.

A discussion of academic environments would not be complete without also acknowledging Tracey, Allen, and Robbins' finding (2012) that environmental constraint is a moderator of the impact of congruence on academic outcomes. In short, they observed that congruence mattered more for "strong" majors that had greater environmental constraint. Weak environments are theorized to be less constraining on behavior, and, therefore, allow for a greater range of personalities to operate successfully in them without the need to change. This concept could be relevant for the study of congruence with both instructional programs and specific classes.

For example, the wide range of occupational outcomes for Philosophy majors may suggest it is a "weak" environment insofar as it is less constraining on occupational choice behavior. Similarly, a class like Music Appreciation may be, out of necessity, a weak environment due to its need to accommodate many students and personality types. In both cases, the role of interest congruence could be lessened due to the decreased need for the individual to conform to environmental demands beyond the application of general academic skills. The same could likely be said for most remedial and general education courses.

If this is the case, data on program choice might be helpful, especially at a PTYC where some students are pursuing more general transfer programs while others are enrolled in more applied career programs. This study did not include major or program choice or any indicators on if a course was for transfer programs, career programs, or both. Thus, there was no way to disaggregate between students based on program of study or identify if a course is predominantly designed for general education or applied learning.

Recommendations. The immediate recommendation that emerges from discussing the limitations related to classifying the academic environment is that there is a desperate need for an updated, comprehensive reclassification of instructional programs using the Holland typology that is publicly available for access or purchase. The ideal would be to avoid the association method altogether and to use either the incumbent, empirical, and/or judgment method. Of these, researchers have already demonstrated a methodology for using the incumbent method with large datasets to develop environmental profile scores for academic majors based on students who have demonstrated success and persistence with their identified majors (Allen & Robbins, 2010; Tracy et al., 2012).

Tracy, Allen, and Robbins (2012) outlined this methodology in a clear and concise way. They used a large sample (n = 21,298) of postsecondary students who had UNIACT results between 1990 and 2003 and had persisted into their sophomore year with at least a 2.00 GPA. This included a wide range of institution types (45 four-year, 38 two-year; 74 public, 9 private).

In addition to complete Holland-type profiles and scores, the researchers also had the last academic major for each student. Environmental profile scores were then developed for programs that had at least 50 representative students by using the mean scores for each RIASEC type across every student within that major.

While it would be a significant undertaking, this could be updated and expanded with an additional decade of data. Even with a higher membership threshold (e.g., 100 or 500 representative students), programs that did not meet with that threshold could be manually reviewed using the judgment method as was done for vocations in the DHOC. Not only would this provide a complete profile with scores for all instructional program, but also this method would also allow for the creation of differentiation scores for each program as Tracy, Allen, and Robbins did (2012). Given their conclusion that the impact of interest-major congruence as a predictor of academic outcomes was moderated by environmental constraint, this would allow researchers to further explore this complicated set of relationships both at the major and classroom level of sub-environment.

A second option would be to use the association method but with a revised methodology that is based on incumbent method principles. Rather than crosswalk *Classification of Instructional Programs* (CIP) with the *Dictionary of Occupational Titles* (DOT), it could be matched with the data the *Occupational Outlook Handbook* (OOH) uses to determine the employment distribution for each of their Field of Study categories (Bureau of Labor Statistics, U.S. Department of Labor, 2023). Rather than associating a major with only the straightforward, most direct careers based on judgment, this method allows for the profile to include the careers that the *actual graduates* are in using that major. While this is not a straightforward incumbent strategy, it allows for the inclusion of data from individuals who actually completed that program. If this method of classifying also allowed for the inclusion of standardized scores for each type, environmental constraint could still be observed.

The final recommendation is to include measures of environmental constraint in analyses of congruence. To what extent does the environment demand certain types of needs from the person? In particular, this question needs to be answered for courses that typically fall within a general education curriculum. This present study would suggest that general education courses do not reward having an interest in their subject material. While the "general" in general education is often conceptualized as referring to a pool of fields that the well-rounded college graduate should have, it may also refer to the structure of the class, regardless of field of study. If so, it makes sense that interest-class subject congruence would be minimized in favor of other types of fit (e.g., ability-requirements, goals-supplies, etc.). A simple way to test this would be to look at either program of choice for the student or program affiliation for the class to see if these variables moderate the relationship between congruence and academic outcomes.

First-Semester GPA

Limitation. The need for consistent environmental data for the sample necessitated that this study focus on a single semester, and the first semester was the obvious choice given that it would offer the largest potential sample size and has been cited as being "especially important" (Bowman et al., 2019, p. 274). While fall-to-fall retention is well researched, it is worth noting that using first-semester GPA as a dependent variable appears to be a somewhat novel approach per the literature. As Gershenfield, Ward Hood, and Zhan (2016) stated, while "there have been many studies involving students' race, income, precollege preparation, and cumulative grade point average (GPA) as predictors of retention…, few take into account first-semester GPA" (p. 469). In their own review of decades of literature, they identified only 15 studies that were relevant and that four of these were unpublished dissertations. It should be emphasized that these are studies that use first-semester GPA as a *predictor* variable. Even fewer use it as a dependent or outcome variable. In the literature review for this study, only one included first-semester GPA as dependent variable (Tracy, Allen, & Robbins, 2012). While not an anticipated addition to the literature, this study highlights some of the limitations that should be taken into consideration when attempting to use this variable.

As the lengthy data cleaning and assumption testing section in Chapter Four demonstrated, first-semester GPA is a particularly unwieldy variable, especially for community college students and especially in analyses that require univariate or multivariate normality. The bunching up of 0.00 and 4.00 GPAs at both ends of the distribution made it difficult to use in a regression analysis. This study used to a combination of Holland's (1997) theory and the literature on college student retention to assert that 0.00 and 4.00 students could be considered distinct subgroups distinguished by having an unknown combination of factors that led to either complete disengagement from or maximum success in their respective academic environments. These subgroups were then removed from the study before proceeding with the analysis. This limited the study from being able to include two groups of students who would certainly be of interest to explore.

Recommendation. Future studies should consider applying alternative statistical analyses that are a better fit for the extreme non-normality of the first-semester GPA variable. In scouring the literature, the only study that seemingly spoke to this particular issue when using first-semester GPA was a dissertation from Saltonstall (2013) in which they initially also set out to use sequential multiple linear regression to predict first-semester GPA at a large university. In discussing their research method, they echoed the same problems encountered in this study and stated that their first-semester GPA variable "was not normally distributed, and that transformation did not result in the data becoming normally distributed" (p. 54-55). In response to this methodological issue, the author changed their statistical analysis altogether to a multinomial logistic regression and recoded their continuous first-semester GPA variable to be categorical for use in this new analysis.

As discussed in Chapter Four, Keith (2015) strongly discourages this practice and asserts that it "throws away" the variable's variance and "reduces its correlation with other variables" (p. 226). However, he does go to acknowledge that it "may make sense to categorize a continuous variable after analysis as an aid in interpretation" or to do so when there are "compelling reasons" (p. 227). While this study did not, one might be able to make the argument that the ordinal nature of GPA combined with the extreme non-normality of first-semester GPA is compelling enough to justify categorization and multinomial logistic regression. This could be a viable path forward not just for studying the impact of congruence but any independent variable of interest on first-semester GPA.

Ill-Structured Problem

Limitations. Braxton, Hirschy, and McClendon's (2004) description of college student departure as being an "ill-structured problem" is a theme that this study returned to throughout its development. Congruence's insignificance aside, the inclusion of age, high school GPA, ethnicity, age, gender, Pell grant eligibility, and parent education level as predictor variables still only accounted for 19.5% of the variance in first-semester GPA and between 7.2% (Cox & Snell R) and 9.6% (Nagelkerke R Square) of the variance when predictor persistence into the second year. These are modest levels of explanatory power.

This is consistent with the results of Allen and Robbins (2010), the only study of congruence reviewed in Chapter Two that included a sample of students from two-year institutions in addition to a sample of students from four-year institutions. These authors used an even more robust set of independent variables including interest-major congruence, motivation, high school GPA, ACT composite score, gender, ethnicity, parent income, and parent education level to predict first-year academic performance and timely degree attainment. While their model was significant for both four- and two-year students, it explained far less of the variation (26%) in the two-year sample than the four-year (50%). They found a similar result when predicting timely degree attainment with the authors reporting that the overall odds ratio "for the 2-year sample was 2.87, substantially less than the 4.11 observed for the 4-year sample" (p. 32). The authors partially attributed these differences to the smaller available sample of two-year students, but also acknowledged that "students who enter 2-year institutions are more likely have life situations that may affect postsecondary success but that are not captured by our data set" (p. 33).

This echoes Tinto's (1993) own skepticism about how his iconic theory could predict two-year student outcomes with the statement that they "likely to experience a wide range of competing external pressures on their time and energies and to be unable to spend significant amounts of time on campus interacting with other students and members of the staff" (p. 78). In short, the reasons that two-year college students do not do well in a given semester or drop out entirely may be even more ill-structured than those at traditional four-year institutions. Like Allen and Robbins (2010), the data set and variables used in this study may not be expansive enough to properly conceptualize the experience of these students.

Of the potential variables that may have been missing, additional considerations from Holland's (1997) theory might be an excellent place to start. Namely, his concepts of consistency and differentiation could be moderating the influence of congruence outcomes. As detailed in Chapter Two consistency refers to the within-profile relationships between the types and the order they are in. Adjacent types are the most consistent, opposite types are the least consistent, and alternate types have intermediate level of consistency. For example, a student with the Holland Code RIC would be considered to have a highly consistent profile since those three types occupy the same half of the hexagon and the secondary and tertiary types are both adjacent to the dominant. On the other hand, a student with the Holland Code RSA would have a less consistent profile since the secondary and tertiary types are opposite and alternate to the dominant respectively.

This is not to say that an individual with a less consistent profile has a deficit. It just suggests that it may be more difficult for them to find an environment that satisfies both areas of interest (e.g., working with both people and things), but it also suggests that they have a wider range of flexibility when it comes to finding environments that could be, at least partially, a good fit for them. It is the person equivalent to the earlier discussed concept of constraint for environments. Highly constraining (i.e., strong) environments have low variation in their responsibilities, and highly consistent individuals have low variation in their interests.

Milsom and Coughlin (2017) used a calculation of consistency as a predictor variable in their study of academic major satisfaction and academic performance and found that it was not significant for either outcome. However, this study has some significant limitations. First, the measure of consistency only took into consideration the dominant and secondary types. Future studies should include or develop methods of calculating consistency that take into consideration, at minimum, the top three if not the full personality profile. Strahan (1987) provides a method for generating a consistency score using the full, three-point Holland Code. Other limitations in the study include self-reported GPA by participants, small sample size (n=99), and lack of any interaction terms in the regression. Thus, the opportunity to explore the concept of consistency and how it impacts the predictive power of congruence is open for empirical analysis.

Holland (1997) defined differentiation as the "magnitude of the difference between highest and lowest scores of the six variables" (p. 57). Differentiation can be visibly observed by looking at the shape of a profile based on the individual scores on each type. For example, look at **Figures 5.1** and **5.2** to see examples of what high versus low differentiation looks like. **Figure 5.1** shows a well-differentiated profile where the dominant, secondary, and tertiary types are clear and distinct from the bottom three, creating high peaks and lower valleys across the profile. **Figure 5.2** is a profile for a student with low differentiation. It is relatively flat, and while there are technically dominant, secondary, and tertiary types, the differences in scores between the dominant and lowest score is quite low (6).

Milsom and Coughlin (2017) also included a measure of differentiation in their analysis, but like consistency, it was not a significant predictor. They calculated differentiation for each student by computing the standard deviation among the six profile scores. Higher standard deviations indicated greater variance and, therefore, greater differentiation in the profile. Lower scores indicate low variance and a flat profile. While this measurement does consider all six scores, it does not distinguish between high, flat profiles versus low, flat profiles. **Figure 5.2** is a relatively low, flat profile which would suggest that the individual indicated modest to low interest in most, if not all, of the tasks that were presented to them in the assessment. **Figure 5.3**

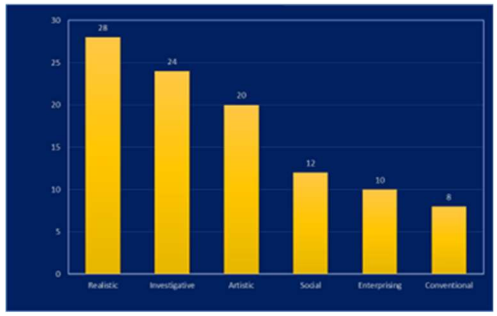

Figure 5.1
Highly Differentiated RIASEC Profile

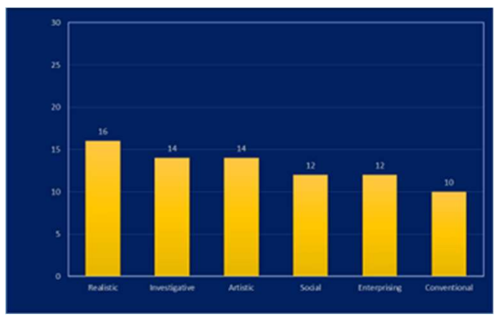

Figure 5.2
Minimally Differentiated RIASEC Profile

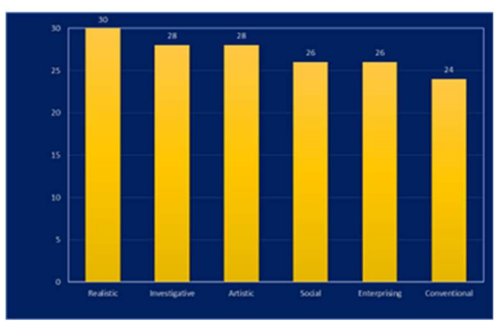

Figure 5.3
High, Minimally Differentiated RIASEC Profile

shows a high, flat profile which indicated the opposite occurrence. Here, the individual has indicated high interest in nearly all of the items presented to them in the assessment. While these two profiles would have similar variance, they represent two very different types of individuals. Thus, a third concept is needed to paint the full picture and may help explain what these authors did not find this to be a significant predictor in addition to the already noted limitations of their study.

The study by Tracy and Robbins (2006) might provide a solution to this issue. They also included a measure that simultaneously accounts for both differentiation and consistency that they called "interest clarity" (p.71). They defined this concept as the "magnitude of differentiation in the entire circular profile" or the "amount of circular variance that exists in a profile" (p. 67). Thus, it is similar to Milsom and Coughlin (2017) in that they use a version of profile variance to determine differentiation, but because it takes the circular structure into account, high but opposing scores (e.g., Realistic and Social) "cancel each other out" (p.71). As they put it, "clarity represents the extent to which an individual has a differentiated and consistent circular profile" (p. 71).

They also accounted for the distinction between high and low flat profiles by including what they called "interest level" or "interest flexibility" as a variable. This is represented as the mean score across all six profile types. As they put it, having "a high mean score indicates a liking for a wide variety of activities" and "a low mean indicates liking relatively few" (p. 67). Including these variables along with moderation interaction terms showed that level of interest moderated the relationship between congruence and persistence. More specifically, "individuals with lower interest levels demonstrated a higher relation of congruence and persistence than individuals with higher interest levels" (P. 85). In other words, congruence may be more important for students with lower overall interest levels. This makes intuitive sense as "these individuals do not have a broad set of likes and as such non-matches in the environment are more crucial for these individuals" (p. 85).

Beyond Holland's (1997), the case should also be made for exploring other types of P-E fit. The point made in the opening chapter that higher education research has focused disproportionately on the person (i.e., student) over the environment still stands. If the conclusion of this study that interest-classroom congruence in the first semester is not a significant predictor of GPA or persistence stands in the face of additional empirical treatment, then it would behoove researchers to study other characteristics of the individual and the extent to which the academic environment aligns with them. Within a few years of Holland publishing the third edition of his book, Tinsley (2000) published a critique of the theory with his article, *The Congruence Myth: An Analysis of the Efficacy of the Person-Environment Fit Model*. In the article, the author reviews the existing literature on the RIASEC typology and congruence, in particular, and concludes that "the evidence demonstrates unequivocally that hexagonal congruence is not a valid predictor of vocational outcomes" and that "[f]urther investigations of the hexagonal structure of the RIASEC dimensions are of little theoretical or practical usefulness" (p. 174 - 175).

While the literature in the last 24 years since this article was published has been kinder to Holland's theory than what Tinsley (2000) advocated, he does insist that P-E fit, models, in general, warrant further ongoing investigation. He wrote that, "The overwhelming conclusion from the research literature is that the P-E fit model provides a valid and useful way of thinking about the interaction between the individual and environment" (p. 150). He also provides a useful distinction between two categories of P-E fit theories that aligns with the Needs-Supplies and Demands-Abilities distinction that Edwards and colleagues (2006) would on go make: 1) the relation of employee desires to job supplies and 2) the relation of employee abilities to job demands. As cited in Chapter Two, there are examples of higher education researchers answering this call (Bohndick, Rosman, Kohlmeyer, & Buhl, 2018; Bowman & Denson, 2014; Etzel & Nagy, 2021; Gilbreath, Kim, & Nichols, 2011; Li, Yao, Chen, & Wang, 2013). However, none of these were done with two-year college students. While this study sought to help remedy this deficit, there is still a great deal of work to be done to move away from what Davidson and Wilson (2017) called the "deficit model" that characterizes research on students at two-year institutions (p. 518).

Recommendations. Per the conceptual underpinnings of Holland's (1997) theory and research and the findings of the aforementioned research, future studies on interest congruence should include measures of complimentary concepts. Namely, the consistency, differentiation, and flexibility of the individual's RIASEC profile should be included in analyses along with interaction terms. For consistency, the literature has provided strategies for using two-point (Milsom & Coughlin, 2017) and three-point (Strahan, 1987) codes, but there is need for development of a more sophisticated consistency measure that considers the full profile.

On the other hand, researchers have demonstrated positive results using existing measures of differentiation and flexibility. Furthermore, these measures are relatively easy to calculate and take into consideration the full profile as long as there are scores available for each RIASEC type. Estimating the variation in a profile using its standard deviation has been used by researchers as a measure of differentiation (Milsom & Coughlin, 2017; Tracy & Robbins, 2006). However, it is important to also control for the final Holland-related concept, flexibility. Interest flexibility or level has been calculated by simply obtaining the mean score across all six types.

This measure has been shown to moderate the relationship between congruence and educational outcomes (Tracy & Robbins, 2006).

Future studies also should not limit themselves to Holland's (1997) theory and interest-environment congruence or the same research methodology. This study adds another voice to the growing chorus of researchers calling for the application of P-E fit theories to higher education research (Davidson & Wilson, 2017; Feldman, Schuetz, 2005; Smart, and Ethington, 2008; Tinsley, 2000). The categorizations of types of fit and fit research outlined by Edwards and colleagues (2006) provided a useful outline for potential avenues of future research. Like Tinsley, they distinguish between two types of fit: 1) needs-supplies fit (N-S fit) and 2) demands-abilities fit (D-A fit). It is worth briefly summarizing these dimensions of fit again to provide context for these recommendations.

"Needs" is an umbrella term for a variety of psychological constructs including desires, values, and goals. It is worth noting that, while some of the previously cited studies have separated individual interests into its own fit dimension, this concept would also fall into the category of needs (Etzel & Nagy, 2021; Li, Yao, Chen, & Wang, 2013). "Supplies" refers to environmental characteristics, reinforcers, payoffs, or benefits that do or do not meet an individual's needs. "Abilities" can include skills, aptitudes, and knowledge. Higher education research traditionally uses high school GPA or ACT/SAT score to operationalize students' abilities upon entering college. The "demands" of the environment have been measured as workload, task requirements, and prerequisite knowledge.

Finally, it worth noting that Edwards and colleagues (2006) also posit a third type of fit, "supplementary fit." They define this as the comparison of the individual's needs or abilities to the other people who make up that environment rather than the environment itself. From the earlier conversation about classifying environments, this would include any methodology that uses the incumbent method to define the characteristics of the environment (Edwards et al., 2006). While the remainder of this section will not include this additional distinction as the nuances and impact of classifying the environment have already been discussed, future researchers may find it helpful as they discuss their P-E fit methodologies.

Edwards and colleagues (2006) identified three types of research methodologies for exploring N-S and D-A fit: 1) atomistic, 2) molecular, and 3) molar. This study is an example of atomistic fit research in that it assessed characteristics of the person and the environment separately and then combined them to represent fit. However, if the members of the sample had completed a survey that asked them to rate their classes on if the class matched their interests enough to stay engaged on a 1-5 Likert scale with -2 (*Not at all*), -1 (*Not really*), 0 (*Just enough*), 1 (*More than enough*), and 2 (*Far more than enough*), this would have been a molecular strategy. Molecular strategies assess the perceived discrepancy between the person and the environment directly and provide a directional relationship (i.e., needs can exceed supplies and supplies can exceed needs). Finally, if the survey instead asked the student to rate their overall fit with the class, this would be a molar approach in that it assesses fit without regard to direction which means there is no indication on if a lack of fit is the result of needs exceeding supplies or vice versa.

Taken together, these concepts provide a broad framework for future higher education researchers to move beyond just assessing student characteristics and evaluate student outcomes in a more holistic and intentional way that considers both the individual and the academic environment. Including molecular and molar measures of interest-classroom fit would have enhanced the analysis in this study. Furthermore, providing other additional indices of fit with which to test the research questions and comparing them could have provided useful insight into the lack of significance of the atomistic measure. Edwards and colleagues (2006) concluded that "atomistic, molecular, and molar approaches to P-E fit should be considered theoretically and empirically distinct" and "are not interchangeable" (p. 822). For example, if Artistic student consistently rated that MUS 1030 as -2, -1, or 0 on a molecular scale, it would provide evidence for the previously stated hypothesis that this classroom environment is weak as it relates to artistic reinforcers.

While somewhat tangential to this study, it is worth also asserting that the P-E fit framework could help future research tie together the sometimes mixed and disparate studies on popular concepts in the field. Take the concept "sense of belonging" and its related concepts (e.g., connectedness, integration, relatedness, membership, etc.), to demonstrate this point. Strayhorn (2012) described this concept as "one term with many meanings" and that "there's a fair amount of disagreement in the literature about the specific definition of sense of belonging" (p. 8). Indeed, the literature often conflates these concepts in defining one another.

For example, McMillian and Chavis (1986) defined "sense of community" as "a feeling that members have of *belonging* [emphasis added] and being important to each other, and a shared faith that their needs will be met by their commitment to be together" (as cited in Mendoza et al., 2016, p. 287). Here, the authors emphasize that "belonging" is one component of the broader concept sense of community. Some applications of McMillan and Chavis's (1986) definition of sense of community have attempted to maintain this distinction (Mendoza et al, 2016). However, Osterman (2000) stated that "the essence" of this definition "is a feeling of belongingness" and go on to use belongingness and sense of community interchangeably (p. 324). This lack of uniformity in how these concepts are used and differentiated from one another makes providing a single definition for sense of belonging, or any of these terms for that matter, difficult.

Strayhorn (2012, p. 3) provided his own attempt at a unifying definition for sense of belonging in higher education settings:

> Sense of belonging refers to students' perceived social support on campus, a feeling or sensation of connectedness, the experience of mattering or feeling cared about, accepted, respected, valued by, and important to the group (e.g., campus community) or others on campus (e.g., faculty, peers). It's a cognitive evaluation that leads to an affective response or behavior.

While helpful, using a P-E fit lens can provide further clarification. Social support is a psychological need and fit occurs when the environment's supplies meet that need. In this case, environmental supplies would be defined as the presence of caring, accepting, respectful, and supportive individuals. When fit between social support needs by the supplies of the individual is achieved, the "sense" or cognitive evaluation of belonging ensues leading to an affective response (e.g., happiness) and behavioral outcomes (e.g., persistence).

This more nuanced, P-E fit informed definition provides a more complete understanding of this outcome and processes that lead to it. Researchers can then choose from atomistic, molecular, and/or molar approaches for measuring social support-environment fit with the understanding that each is theoretically distinct and should be used to inform one another rather than used interchangeable. Researchers could then explore if other types of fit predict sense of belonging. For example, perhaps it is not just social support that leads a student to feel like they belong on a college campus? One could see a situation where, despite an environment full of caring, supportive individuals, a student departs from that environment because they perceive that their skills do not match the demands or their goals do not match the opportunities. This is just one well-researched concept among many in higher education that could be further organized and clarified by using a P-E fit theoretical framework.

Conclusion

The goal of this study was to better understand the contributing factors to first-year student performance and persistence at a large, public community college by applying Holland's (1997) concept of congruence to determine if it significantly predicted these outcomes over and above high school performance and demographic characteristics. By using this theoretical framework and asking these research questions, multiple identified gaps in the literature could be partially addressed simultaneously. This included adding to the sparse research on students at public two-year colleges (Davidson & Wilson, 2017), applying a P-E fit theoretical framework (Schuetz, 2005), and applying Holland's theory, specifically, to higher education outcomes (Smart et al., 2000). The results of the analyses indicated that interest-classroom congruence in the first semester is not a significant predictor of first-semester GPA or persistence in the second year. Despite these results, it is the contention of this author that the book should be far from closed on the topic.

When this study was originally conceived, it was done so from the perspective that it would be the next in a long lineage of established research. Indeed, the theoretical framework and variables used in this study are well-researched topics in a general sense. However, after developing the methodology, conducting the analyses, and reflecting on the results, it is only *post hoc* that it has become clear that this study is more exploratory than initially realized. Swedberg (2020) states that one of the reasons that an exploratory study is conducted is to "to generate new and interesting hypotheses about a topic that is already known" (p. 28). He goes on to further elaborate on what he calls an “informal exploratory study” or “pre-study” that helps the researcher “know something about her topic before the design for the main study has been drawn up” and “come up with new ideas” (p. 33).

While it was not the goal to relegate the work done in this dissertation to a pre-study, the process of implementing this study has generated new ideas and identified gaps in the literature that warrant further exploration. This includes measuring needs/abilities-classroom fit, explaining the variance in two-year student departure, studying first-semester GPA, developing an updated and more sophisticated RIASEC classification of instructional programs, and the potential of using P-E fit theories as a framework for higher education research.

List of References

ACT. (1995). *Technical manual: Revised unisex edition of the ACT interest inventory*. ACT, Inc.

ACT. (2009). *ACT interest inventory technical manual.* ACT, Inc. Retrieved on October 1, 2023, from https://www.act.org/content/act/en/research/reports/technical-manuals-and-fairness-reports.html

Adams, P., Adams, W., Franklin, D., Gulick, D., Gulick, F., Shearn, E.,… Mireles, S. V. (2012). Remediation: Higher education's bridge to nowhere. Retrieved December 28, 2018, from www.insidehighered.com

Allen, J., & Robbins, S. B. (2008). Prediction of college major persistence based on vocational interests, academic preparation, and first-year academic performance. *Research in Higher Education*, *49*(1), 62–79. https://doi.org/10.1007/s11162-007-9064-5

Allen, J., & Robbins, S. B. (2010). Effects of interest–major congruence, motivation, and academic performance on timely degree attainment. *Journal of Counseling Psychology*, 57, 23–35. https://doi.org/10.1037/a0017267

Astin, A. W. & Holland, J. L. (1961). The Environmental Assessment Technique: A way to measure college environments. *Journal of educational psychology* 52.6 (1961): 308–316. https://doi.org/10.1037/h0040137

Astin, A. W. (1965). Classroom environment in different fields of study. *Journal of Educational Psychology*, *56*(5), 275–282. https://doi.org/10.1037/h0022514

Bailey, T., Jenkins, J., & Leinbach, T. (2005). What we know about community college low-income and minority student outcomes: Descriptive statistics from national surveys. *Community College Research Center*. Retrieved on January 3, 2019, from https://ccrc.tc.columbia.edu/media/k2/attachments/low-income-minority-completion.pdf

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.

Bohndick, C, Rosman, T., Kohlmeyer, S., & Buhl, H. (2018). The interplay between subjective abilities and subjective demands and its relationship with academic success: An application of the person-environment fit theory. *Higher Education*, *75*(5), 839–854. https://doi.org/10.1007/s10734-017-0173-6

Bowman, N. A., & Denson, A. (2014). A missing piece of the departure puzzle: Student–institution fit and intent to persist. *Research in Higher Education, 55*(2), 123-142. https://doi.org/10.1007/s11162-013-9320-9

Bowman, N. A., Jarratt, L., Jang, N., & Bono, T. J. (2019). The unfolding of student adjustment during the first semester of college. *Research in Higher Education*, *60*(3), 273–292. https://doi.org/10.1007/s11162-018-9535-x

Braxton, J. M., Hirschy, A. S., & McClendon, S. A. (2004). *Understanding and reducing college student departure* (ASHE-ERIC higher education report, v. 30, no. 3). Jossey-Bass.

Brown, S. D., & Gore. P. A. (1994). An evaluation of interest congruence indices: Distribution characteristics and measurement properties. *Journal of Vocational Behavior*, *45*(3), 310-327. https://doi.org/10.1006/jvbe.1994.1038

Bui, K. V. (2002). First-generation college students at a four-year university: Background characteristics, reasons for pursuing higher education, and first-year experiences. *College Student Journal*, *36*(1), 3.

Bureau of Labor Statistics, U.S. Department of Labor (2023). Field of degree: Philosophy and religion. *Occupational Outlook Handbook*. Retrieved on October 14, 2023 from https://www.bls.gov/ooh/field-of-degree/philosophy-and-religion/philosophy-and-religion-field-of-degree.htm.

Camp, C. C. & Chartrand, J. M. (1992). A comparison and evaluation of interest congruence indices. *Journal of Vocational Behavior, 41*(2), 162-182. https://doi.org/10.1016/0001-8791(92)90018-U

Cataldi, E., Bennet, C., & Chen, X. (2018). First-generation students: College access, persistence, and postbachelor's outcomes. U.S. Department of Education, National Center for Education Statistics. Retrieved January 15, 2019, from https://nces.ed.gov/pubs2018/2018421.pdf

Chen, X. (2016). Remedial coursetaking at U.S. public 2- and 4-year institutions: Scope, experiences, and outcomes (NCES 2016-405). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved on January 8. 2019, from http://nces.ed.gov/pubsearch

Cohen, A. M., Brawer, F. B., & Kisker, C. B. (2013). *The American community college* (6th ed.). Jossey-Bass.

Complete College America. (n.d.). Retrieved March 21, 2021, from https://completecollege.org/our-work/

Davidson, J. C., & Wilson, K. B. (2017). Community college student dropouts from higher education: toward a comprehensive conceptual model. *Community College Journal of Research and Practice, 41*(8), 517-530. https://doi.org/10.1080/10668926.2016.1206490

Deil-Amen, R. (2011). Socio-academic integrative moments: Rethinking academic and social integration among two-year college students in career-related programs. *The Journal of Higher Education (Columbus)*, *82*(1), 54–91. https://doi.org/10.1080/00221546.2011.11779085

Edwards, J. R., Cable, D. M., Williamson, I. O., Lambert, L. S., & Shipp, A. J. (2006). The Phenomenology of Fit: Linking the Person and Environment to the Subjective Experience of Person-Environment Fit. *Journal of Applied Psychology*, *91*(4), 802–827. https://doi.org/10.1037/0021-9010.91.4.802

Etzel, J. M., & Nagy, G. (2021). Challenging the multidimensional conception of perceived Person-Environment Fit: Are specific fit dimensions related to educational outcomes beyond a higher-order factor? *European Journal of Psychological Assessment: Official Organ of the European Association of Psychological Assessment*, *37*(5), 368–376. https://doi.org/10.1027/1015-5759/a000622

Feldman, K. A., Smart, J. C., & Ethington, C. A. (2004). What do college students have to lose? Exploring the outcomes of differences in person-environment fits. *The Journal of Higher Education*, *75*(5), 528-555. https://doi.org/10.1080/00221546.2004.11772336

Feldman, K. A., Ethington, C. A., & Smart, J. C. (2008). Using Holland's theory to study patterns of college student success: The impact of major fields on students. In *Higher Education: Handbook of Theory and Research* (Vol. 23, Higher Education: Handbook of Theory and Research, pp. 329-380). Springer Netherlands.

Ford, C. (2018). Interpreting log transformation in a linear model. University of Virgina Library. Retrieved August 8, 2023, from https://data.library.virginia.edu/interpreting-log-transformations-in-a-linear-model/

Gati, I. (1985). Description of alternative measures of the concepts of vocational interest: Crystallization, congruence, and coherence. *Journal of Vocational Behavior, 27*(1), 37-55. https://doi.org/10.1016/0001-8791(85)90051-X

Gershenfeld, S., Ward Hood, D., & Zhan, M. (2016). The role of first-semester GPA in predicting graduation rates of underrepresented students. *Journal of College Student Retention: Research, Theory & Practice*, *17*(4), 469–488. https://doi.org/10.1177/1521025115579251

Gilbreath, B., Kim, T., & Nichols, B. (2011). Person-environment fit and its effects on university students: A response surface methodology study. *Research in Higher Education, 52*(1), 47-62. https://doi.org/10.1007/s11162-010-9182-3

Ginder, S. A., Kelly-Reid, J. E., and Mann, F. B. (2018). Postsecondary institutions and cost of attendance in 2017– 18; Degrees and other awards conferred, 2016–17; and 12-month enrollment, 2016–17. U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved June 25, 2019, from http://nces.ed.gov/pubsearch

Goodman, A. (2024). Post-traditional students' perceptions of mattering: The role of faculty and student interaction. *The Journal of Continuing Higher Education*, *72*(1), 17–32. https://doi.org/10.1080/07377363.2022.2108642

Gottfredson, G. D., Holland, J. L., & Ogawa, D. K. (1982). *Dictionary of Holland occupational codes*. Consulting Psychologists Press.

Gottfredson, G. D., & Holland, J. L. (1989). *Dictionary of Holland occupational codes* (2nd ed.). Psychological Assessment Resources, Inc.

Gottfredson, G. D., & Holland, J. L. (1996). *Dictionary of Holland occupational codes* (3rd ed.). Psychological Assessment Resources, Inc.

Gottfredson, L. S. (1981). Circumscription and compromise: A developmental theory of occupational aspirations. *Journal of Counseling Psychology, 28*(6), 545-579. https://doi.org/10.1037/0022-0167.28.6.545

Gottfredson, L. S. & Lapan, R. T. (1997). Assessing gender-based circumscription of occupational aspiration. *Journal of Career Assessment*, *5*(4), 419-41.

Governor Bill Haslam Launches Drive to 55 Initiative. (2013). *UT Advocacy*. Retrieved October 8, 2018, from https://advocacy.tennessee.edu/2013/09/04/drive-to-55-initiative/

Grimes, S. K., & David, K. C. (1999). Underprepared community college students: Implications of attitudinal and experiential differences. *Community College Review*, *27*(2), 73-92. https://doi.org/10.1177/009155219902700204

Hackett, G. & Betz, N. E. (1981). A self-efficacy approach to the career development of women. *Journal of Vocational Behavior*, *18*(3), 326–339. https://doi.org/10.1016/0001-8791(81)90019-1

Harmon, L., Hansen, J., Borgen, F., & Hammer A. (1994). *Strong interest inventory: Applications and technical guide*. Stanford University Press.

Healy, C. C. & Mourton, D. L. (1983). Derivatives of the Self-Directed Search: Potential clinical and evaluative uses. *Journal of Vocational Behavior, 23*(3), 318-28. https://doi.org/10.1016/0001-8791(83)90045-3

Hirschi, A., & Läge, D. (2007). Holland’s secondary constructs of vocational interests and career choice readiness of secondary students. *Journal of Individual Differences*, *28*, 205–218. https://doi.org/10.1027/1614-0001.28.4.205

Hoachlander, G., Sikora, A., & Horn, L. (2003). Community college students: Goals, academic preparation, and outcomes. U.S. Department of Education, National Center for Education Statistics.

Holland, J. L. (1963). Explorations of a theory of vocational choice and achievement: A four-year prediction study. *Psychological Reports*, 12, 547-594. https://doi.org/10.2466/pr0.1963.12.2.547

Holland, J. L. (1973). *Making vocational choices: A theory of careers*. Prentice-Hall.

Holland, J. L. (1994). *The self-directed search: Professional manual—Form R*. Psychological Assessment Resources.

Holland, J. L., Fritzsche, B. A., & Powell, A. B. (1994). *Self-directed search technical user's guide*. Psychological Assessment Resources.

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Psychological Assessment Resources.

Hurtado, S., & Carter, D. (1997). Effects of college transition on perceptions of the campus racial climate on students' sense of belonging. *Sociology of Education*, *70*, 324-345.

Iachan, R. (1984). A measure of agreement for use with the Holland classification system. *Journal of Vocational Behavior, 24*(2), 133-41. https://doi.org/10.1016/0001-8791(84)90001-0

Inman, W. E., & Mayes, L. (1999). The importance of being first: Unique characteristics of first generation community college students. *Community College Review, 26*(4), 3–22. https://doi.org/10.1177/009155219902600402

Ishitani, T. T. (2003). A longitudinal approach to assessing attrition behavior among first-generation students: Time-varying effects of pre-college characteristics. *Research in Higher Education, 44*(4), 433-449. https://doi.org/10.1023/A:1024284932709

Iversen, & Norpoth, H. (1987). *Analysis of variance* (2nd ed.). SAGE.

Karp, M. M., Hughes, K. L., & O'Gara, L. (2010). An exploration of Tinto's integration framework for community college students. *Journal of College Student Retention, 12*(1), 69-86. https://doi.org/10.2190/CS.12.1.e

Kast, M. (2019). New program will help Knox County students attend and finish college. *Knox News.* Retrieved June 25, 2019, from https://www.knoxnews.com/story/news/education/2019/06/19/new-program-help-knox-county-students-attend-and-finish-college/1489776001/

Keith, T. Z. (2015). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling*. (2nd ed.). Taylor & Francis.

Kline, R. B. (2010). *Principles and practice of structural equation modeling*. (3rd ed.). Guilford Press.

Kubala, K. & Borglum, T. (2000). Academic and social integration of community college students: A case study. *Community College Journal of Research and Practice, 24*(7), 567–576. https://doi.org/10.1080/10668920050139712

Kwak, J. C. & Pulvino, C. J. (1982). A mathematical model for comparing Holland's personality and environmental codes. *Journal of Vocational Behavior, 21*(2), 231-241. https://doi.org/10.1016/0001-8791(82)90032-X

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, 45(1), 79-122. https://doi.org/10.1006/jvbe.1994.1027

Lewis, P., & Rivkin, D. (1999). Development of the O*NET interest profiler. National Center for O*NET Development, Raleigh, NC. Retrieved December 27, 2018, from https://www.onetcenter.org/reports/IP.html

Li, Y., Yao, X., Chen, K., & Wang, Y. (2013). Different fit perceptions in an academic environment: Attitudinal and behavioral outcomes. *Journal of Career Assessment*, *21*(2), 163–174. https://doi.org/10.1177/1069072712466713

Luke, C., Redekop, F., & Burgin, C. (2014). Psychological factors in community college student retention. *Community College Journal of Research and Practice, 39*(3), 1-13. https://doi.org/10.1080/10668926.2013.803940

Mayers, A. (2013). *Introduction to statistics and SPSS in psychology*. Pearson.

Mayo, M. M., & Christenfeld, N. (1999). Gender, race, and performance expectations of college students. *Journal of Multicultural Counseling and Development, 27*(2), 93-104. https://doi.org/10.1002/j.2161-1912.1999.tb00217.x

McFarland, J., Hussar, B., Wang, X., Zhang, J., Wang, K., Rathbun, A., Barmer, A., Forrest Cataldi, E., and Bullock Mann, F. (2018). The condition of education 2018. U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved August 25, 2018, from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018144

Mehta, S. S., Newbold, J. J., & O'Rourke, M. A. (2011). Why do first-generation students fail? *College Student Journal, 45*(1), 20-35.

Mendoza, P., Suarez, J., & Bustamante, E. (2016). Sense of community in student retention at a tertiary technical institution in Bogotá: An ecological approach. *Community College Review, 44*(4), 286-314.

Miller, M. J., Wells, D., Springer, T. P., & Cowger, E. J. (2003). Do types influence types? Examining the relationship between students' and parents' Holland codes. *College Student Journal, 37*(2), 190.

Milsom, A. & Coughlin, J. (2017). Examining person-environment fit and academic major satisfaction. *Journal of College Counseling, 20*(3), 250-262. https://doi.org/10.1002/jocc.12073

Mishory, J., & Granville, P. (2019). Policy design matters for rising "free college" aid. *The Century Foundation*. Retrieved June 25, 2019, from https://tcf.org/content/commentary/policy-design-matters-rising-free-college-aid/?session=1&agreed=1

Monks, J. (2000). The returns to individual and college characteristics: Evidence from the National Longitudinal Survey of Youth. *Consortium on Financing Higher Education, 19*, 279-289.

Morrow, J. A., & Skolits, G. J. (2017). Twelve steps of quantitative data cleaning: Strategies for dealing with dirty evaluation data. Presented at the annual conference of the American Evaluation Association, November, 2017.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*, 30–38. https://doi.org/10.1037/0022-0167.38.1.30

Nakajima, M. A., Dembo, M. H., & Mossler, R. (2012). Student persistence in community colleges. *Community College Journal of Research and Practice, 36*(8), 591-613. https://doi.org/10.1080/10668920903054931

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2019). Beginning postsecondary longitudinal study (BPS:12/17). Retrieved on September 27, 2023, from https://nces.ed.gov/datalab/powerstats/71-beginning-postsecondary-students-2012-2017

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2018). 2015–16 National Postsecondary Student Aid Study (NPSAS:16). Retrieved on September 27, 2023, from https://nces.ed.gov/datalab/powerstats/121-national-postsecondary-student-aid-study-2016-undergraduates

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (1991). A classification of instructional programs (1990 ed.). Washington, DC: U.S. Government Printing Office.

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2000). Classification of instructional programs (2000 ed.). Washington, DC: U.S. Government Printing Office.

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2015). Postsecondary attainment: Differences by socioeconomic status. Retrieved on January 3, 2019, from https://nces.ed.gov/programs/coe/pdf/coe_tva.pdf

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2019). Undergraduate retention and graduation rates. Retrieved November 9, 2019, from https://nces.ed.gov/programs/coe/indicator_ctr.asp

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2020). Classification of instructional programs (2020 ed.). Washington, DC: U.S. Government Printing Office.

Nauta, M. M. (2007). Assessing college students' satisfaction with their academic majors. *Journal of Career Assessment*, *15*(4), 446–462. https://doi.org/10.1177/1069072707305762

Nauta, M. M. (2013). Holland's theory of vocational choice and adjustment. In S. D. Brown & R. W. Lent (Eds.), *Career development and counseling: Putting theory and research into work* (pp. 55-82). John Wiley & Sons.

NSC Research Center. (2018). Persistence & retention – 2018. National Student Clearinghouse Research Center. Retrieved January 3, 2019, from https://nscresearchcenter.org/snapshotreport33-first-year-persistence-and-retention/

Nye, C. D., Su, R., Rounds, J., & Drasgow (2012). Vocational interests and performance: A quantitative summary of over 60 years of research. *Perspectives on Psychological Science, 7*, 384–403. https://doi.org/10.1177/1745691612449021

Osterman, K. F. (2000). Students' need for belonging in the school community. *Review of Educational Research*, 70(3), 323–367.

Pascarella, E. T., & Chapman, D. W. (1983). A multiinstitutional, path analytic validation of Tinto's model of college withdrawal. *American Educational Research Journal*, *20*(1), 87–102. https://doi.org/10.2307/1162676

Pascarella, E. T., Smart, J. C., & Ethington, C. A. (1986). Long-term persistence of two-year college students. *Research in Higher Education*, *24*(1), 47–71. https://doi.org/10.1007/BF00973742

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of research*. San Francisco: Jossey-Bass.

Pascarella, E. T., Terenzini, P. T., & Wolfle, L. M. (1986). Orientation to college and freshman year persistence/withdrawal decisions. *The Journal of Higher Education, 57*(2), 155-175. https://doi.org/10.2307/1981479

Perna, L. W., & Titus, M. A. (2005). The relationship between parental involvement as social capital and college enrollment: An examination of racial/ethnic group differences. *The Journal of Higher Education, 76*(5), 485-518. https://doi.org/10.1080/00221546.2005.11772296

Person-Environment Fit. (2004). *Encyclopedia of Applied Psychology*.

Porter, S. R. & Umbach, P. D. (2006). College major choice: An analysis of person-environment fit. *Research in Higher Education, 47*(4), 429–449. https://doi.org/10.1007/s11162-005-9002-3

Purdue University Career Success. (n.d.). [Image of Holland Hexagon]. Retrieved from https://www.cco.purdue.edu/students/hollandhexagon

Purpose first. (2020). Retrieved March 21, 2021 from https://completecollege.org/purposefirst/

Robbins, P. I., Thomas, L. E., Harvey, D. W., & Kandefer, C. (1978). Career change and congruence of personality type: An examination of DOT-derived work environment designations. *Journal of Vocational Behavior, 13*(1), 15-25. https://doi.org/10.1016/0001-8791(78)90067-2

Rocconi, L. M., Liu, X., & Pike, G. R. (2020). The impact of person-environment fit on grades, perceived gains, and satisfaction: an application of Holland's theory. *Higher Education*, *80*(5), 857–874. https://doi.org/10.1007/s10734-020-00519-0

Rounds, J., Smith, T., Hubert, L., Lewis, P., and Rivkin, D. (1999). Development of occupational interest profiles for O*Net. Retrieved December 1, 2020, from https://www.onetcenter.org/dl_files/OIP.pdf

Rounds, J., Su, R., Lewis, P., & Rivkin, D. (2010). O*NET Interest Profiler Short form psychometric characteristics: Summary. National Center for O*NET Development.

Retrieved December 27, 2018, from http://www.onetcenter.org/reports/IPSF_Psychometric.html

Rounds, J., Wee, C. J., Cao, M., Song, C., & Lewis, P. (2016). Development of an O*NET® Mini Interest Profiler (Mini-IP) for mobile devices: Psychometric characteristics. National Center for O*NET Development. Retrieved December 27, 2018, from: https://www.onetcenter.org/reports/Mini-IP.html

Rounds, J., Hoff, K., Chu, C., Lewis, P., Gregory, C. (2018). O*NET® interest profiler short form paper-and-pencil version: evaluation of self-scoring and psychometric characteristics. National Center for O*NET Development. Retrieved December 27, 2018, from https://www.onetcenter.org/dl_files/IPSF_PP.pdf

Saltonstall, M. (2013). *Predicting college success: Achievement, demographic, and psychosocial predictors of first-semester college grade point average*. ProQuest Dissertations Publishing.

Schuetz, P. (2005). UCLA community college review: Campus environment: A missing link in studies of community college attrition. *Community College Review, 32*(4), 60–80. https://doi.org/10.1177/009155210503200405

Skomsvold, P. (2014). Profile of undergraduate students: 2011–12 (NCES 2015-167). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Shapiro, D., Dundar, A., Huie, F., Wakhungu, P., Yuan, X., Nathan, A., & Hwang, Y. (2017). A national view of student attainment rates by race and ethnicity – Fall 2010 cohort (Signature Report No. 12b). Herndon, VA: National Student Clearinghouse Research

Center. Retrieved August 26, 2018, from https://nscresearchcenter.org/wp-content/uploads/Signature12-RaceEthnicity.pdf

Smart, J. C., Feldman, K. A, & Ethington, C. A. (2000). *Academic disciplines: Holland's theory and the study of college students and faculty.* Vanderbilt University Press.

Spencer, G., de Novais, J., Chen-Bendle, E., & Ndika, E. (2024). A dream deferred: Post-traditional college trajectories and the evolving logic of college plans. *The Journal of Higher Education (Columbus)*, 1–23. https://doi.org/10.1080/00221546.2023.2216611

Spokane, A. R. (1985). A review of research on person-environment congruence in Holland's theory of careers. *Journal of Vocational Behavior, 26*(3), 306-43. https://doi.org/10.1016/0001-8791(85)90009-0

Spokane, A. R., Meir, E. I., & Catalano, M. (2000). Person–environment congruence and Holland's theory: A review and reconsideration. *Journal of Vocational Behavior, 57*(2), 137-187. https://doi.org/10.1006/jvbe.2000.1771

Strahan, R. F. (1987). Measures of consistency for Holland-type codes. *Journal of Vocational Behavior*, *31*(1), 37–44. https://doi.org/10.1016/0001-8791(87)90033-9

Strayhorn, T. L. (2012). *College students' sense of belonging: A key to educational success for all students*. Routledge.

Strong Interest Inventory. (2004). *The Concise Corsini Encyclopedia of Psychology and Behavioral Science.*

Stuart, G., Rios-Aguilar, C., & Deil-Amen, R. (2014). How much economic value does my credential have?: Reformulating Tinto's model to study students' persistence in community colleges. *Community College Review, 42*(4), 327–341. https://doi.org/10.1177/0091552114532519

Swedberg, R. (2020). Exploratory Research. In *The Production of Knowledge* (pp. 17–41). Cambridge University Press. https://doi.org/10.1017/9781108762519.002

Durkheim, E. (1951). *Suicide*. Free Press.

Tabachnick, B. G., & Fidell, L. S. (2018). *Using multivariate statistics*. (7th ed.). Pearson.

Tennessee Reconnect. (n.d.). *Tennessee Reconnect grant*. Tennessee Reconnect. https://tnreconnect.gov/Tennessee-Reconnect-Grant

Terenzini, P. T., Springer, L, Yaeger, P. M., Pascarella, E. T., & Nora, A. (1996). First-generation college students: Characteristics, experiences, and cognitive development. *Research in Higher Education, 37*(1), 1-22. https://doi.org/10.1007/BF01680039

The National Center of Public Policy and Higher Education. (2011). Affordability and transfer: Critical to increasing baccalaureate degree completion. Retrieved August 26, 2018, from http://www.highereducation.org/reports/pa_at/PolicyAlert_06-2011.pdf

Tinsley, H. E. A. (2000). The Congruence Myth: An Analysis of the Efficacy of the Person–Environment Fit Model. *Journal of Vocational Behavior*, *56*(2), 147–179. https://doi.org/10.1006/jvbe.1999.1727

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research* 45: 89-125. https://doi.org/10.3102/00346543045001089

Tinto, V. (1986). *Leaving college: Rethinking the causes and cures of student attrition*. University of Chicago Press.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). University of Chicago Press.

tnAchieves. (n.d.). *Frequently asked questions*. tnAchieves. Retrieved November 20, 2024, from https://www.tnachieves.org/faq

Tracey, T., & Robbins, S. B. (2006). The interest–major congruence and college success relation: A longitudinal study. *Journal of Vocational Behavior, 69*(1), 64-89. https://doi.org/10.1016/j.jvb.2005.11.003

Tracey, T., Allen, J., & Robbins, S. (2012). Moderation of the relation between person-environment congruence and academic success: Environmental constraint, personal flexibility and method. *Journal of Vocational Behavior*, *80*(1), 38–49. https://doi.org/10.1016/j.jvb.2011.03.005

Wall, J. E., & Baker, H. E. (1997). The Interest-Finder: Evidence of validity. *Journal of Career Assessment*, 5, 255-273. https://doi.org/10.1177/106907279700500301

Wolfe, L. K., & Betz, N. E. (1981). Traditionality of choice and sex-role identification as moderators of the congruence of occupational choice in college women. *Journal of Vocational Behavior, 18*(1), 43-55. https://doi.org/10.1016/0001-8791(81)90028-2

Voorhees, R. (1987). Toward building models of community college persistence: A logit analysis. *Research in Higher Education, 26*(2), 115–129. https://doi.org/10.1007/BF00992024

Xu, H. (2023). Vocational interests: Conceptual issues, research findings, and practical implications. In *Career Psychology* (pp. 145–167). American Psychological Association. https://doi.org/10.1037/0000339-008

Zener, T. B. & Schnuelle, L. (1976). Effects of the self-directed search on high school students. *Journal of Counseling Psychology, 23*(4), 353-359. https://doi.org/10.1037/0022-0167.23.4.353

Appendices

Appendix A

O*Net Interest Profiler Short Form Paper-and-Pencil Version

O*NET INTEREST PROFILER SHORT FORM

Read the 60 work activities below. Place a check in the box by the activities you would like to do. **Do not** think about how much education/training is needed or how much money you will make! Count the number of checks for each shaded section and write that total in the box to the right of each section. These are your scores for each interest area.

☐ Build kitchen cabinets	☐ Drive a truck to deliver packages to offices and homes	Total
☐ Lay brick or tile	☐ Test the quality of parts before shipment	
☐ Repair household appliances	☐ Repair and install locks	
☐ Raise fish in a fish hatchery	☐ Set up and operate machines to make products	
☐ Assemble electronic parts	☐ Put out forest fires	
	Realistic checks =	☐
☐ Develop a new medicine	☐ Investigate the cause of a fire	Total
☐ Study ways to reduce water pollution	☐ Develop a way to better predict the weather	
☐ Conduct chemical experiments	☐ Work in a biology lab	
☐ Study the movement of planets	☐ Invent a replacement for sugar	
☐ Examine blood samples using a microscope	☐ Do laboratory tests to identify diseases	
	Investigative checks =	☐
☐ Write books or plays	☐ Paint sets for plays	Total
☐ Play a musical instrument	☐ Write scripts for movies or television shows	
☐ Compose or arrange music	☐ Perform jazz or tap dance	
☐ Draw pictures	☐ Sing in a band	
☐ Create special effects for movies	☐ Edit movies	
	Artistic checks =	☐
☐ Teach an individual an exercise routine	☐ Teach children how to play sports	Total
☐ Help people with personal or emotional problems	☐ Teach sign language to people who are deaf or hard of hearing	
☐ Give career guidance to people	☐ Help conduct a group therapy session	
☐ Perform rehabilitation therapy	☐ Take care of children at a day-care center	
☐ Do volunteer work at a non-profit organization	☐ Teach a high-school class	
	Social checks =	☐
☐ Buy and sell stocks and bonds	☐ Negotiate business contracts	Total
☐ Manage a retail store	☐ Represent a client in a lawsuit	
☐ Operate a beauty salon or barber shop	☐ Market a new line of clothing	
☐ Manage a department within a large company	☐ Sell merchandise at a department store	
☐ Start your own business	☐ Manage a clothing store	
	Enterprising checks =	☐
☐ Develop a spreadsheet using computer software	☐ Calculate the wages of employees	Total
☐ Proofread records or forms	☐ Inventory supplies using a hand-held computer	
☐ Install software across computers on a large network	☐ Record rent payments	
☐ Operate a calculator	☐ Keep inventory records	
☐ Keep shipping and receiving records	☐ Stamp, sort, and distribute mail for an organization	
	Conventional checks =	☐

In the boxes below, write the names of the interest areas with the three highest scores. The first box is your highest or primary interest. If there are ties, choose the interest with activities that you think are the best fit for you.

1 ☐ 2 ☐ 3 ☐

Sponsored by the U.S. Department of Labor, Employment & Training Administration. Developed by the National Center for O*NET Development (v1)

Note. Adapted from Rounds, Hoff, Chu, Lewis, and Gregory (2018)

Appendix B

Pellissippi State Community College IRB Approval Letter

PELLISSIPPI STATE
COMMUNITY COLLEGE

INSTITUTIONAL EFFECTIVENESS, ASSESSMENT AND PLANNING

September 13, 2021

Christian Lockhart
5501 Lakeshore Drive
Knoxville, TN 37920

Dear Mr. Lockhart,

The Institutional Review Board at Pellissippi State Community College has received your application for permission to conduct your study, *The Relationship between Person-Environment Congruence and Academic Success at a Community College.* The Board believes the design of your study meets the Federal requirements for protection of human participants. Your application has received approval as required by PSCC Policy 08:02:01 Conducting Research at Pellissippi State.

Any significant changes in this research project must be reviewed by the IRB at Pellissippi State. All information must be received and recorded in such a manner that subjects cannot be identified, directly or indirectly through identifiers linked to the participants. Please submit any changes in writing. The College looks forward to seeing the results of the study.

Sincerely,

Nancy A. Ramsey

Nancy A. Ramsey, Chair
Institutional Review Board

VITA

Christian Lockhart grew up in East Tennessee and currently resides in Knoxville, TN. He graduated from Maryville College in 2007 with a Bachelor of Arts in Philosophy. In 2013, he earned a Master's degree in Mental Health Counseling from Lincoln Memorial University. During his graduate studies, Christian worked at residential treatment facilities for adolescents with alcohol, drug, and behavioral challenges. Christian began his career in higher education as a counseling intern at Pellissippi State Community College, where he later served as a part-time Academic Advisor before transitioning to a full-time advising position in 2014. In 2017, he joined Pellissippi State's Counseling Services department and became a Licensed Professional Counselor in the state of Tennessee. In 2023, Christian assumed his current role as the Director of Student Care and Advocacy, where he leads efforts to support students facing challenges related to basic needs. He also does private practice counseling in the community.