A Mixed-Methods Theory-Based Evaluation of a Program Supporting Underrepresented Minority STEM Students

A DISSERTATION

SUBMITTED TO THE FACULTY OF

THE UNIVERSITY OF MINNESOTA

BY

Yue Zhang

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Advisor: Dr. Stuart Yeh

December 2022

© Yue Zhang, 2022

Acknowledgements

It's been almost ten years since my first day in Minnesota. In all these years, I have made many choices, large and small. None of any other choices and decisions has required more deliberation and been more life-changing than entering the Ph.D. program, which took me five and a half years to complete. Approaching the end of this journey, I cannot help but look back at the hardships and challenges along the way. I experienced the long separation from my beloved, the physical and mental struggle with the unexpected illness, and the isolation brought about by the worldwide pandemic. I once deeply felt the smallness and loneliness of being one life. It would be disingenuous to say that I never thought of giving up, but I did not—because of all the help, support, encouragement, accompany, and inspiration from all the great people who have always been by my side.

Among these people, my advisor, Dr. Stuart Yeh, is undoubtedly the one in the front. I have benefited greatly from his mentoring. He has been the most responsible, approachable, and trustworthy person who shows and guides my way. I appreciate him for always understanding my choices and helping me to carve my rough thoughts into academic work that stands up to close examination. His insistence on seeking truth from facts and his spirit of rigorous scholarship has fostered the same pursuits in me.

I thank my committee members, Dr. David Johnson, Dr. Andrew Zieffler, Dr. Joshua Collins, and Dr. Michael Stebleton, for all their intelligence, thoughtful suggestions, and guidance on my study. I have enjoyed discussions with them during exams and in their courses. Their expertise in theories, methodologies, and subject matters has greatly benefitted me.

I thank my supervisors at the North Star STEM Alliance (Dr. Keisha Varma, Simone Gbolo, and *the late* Anne Hornickel), the University of Minnesota Extension (Dr. Emily Becher and Dr. Mary Jo Katras), the Center for Urban and Regional Affairs (Dr. Brittany Lewis and Dr. Shana Riddick), and other college departments (Dr. Joyce Serido, Dr. John LaVelle, Dr. Frank Liu, and Dr. Jennifer van de Ligt) where I used to hold graduate assistantships. They provided me with learning opportunities through hands-on experiences; they empowered me with freedom; they trusted in my knowledge and skills and were never stingy with their praise and encouragement. Importantly, they financially supported me so that I could complete my studies with a roof over my head.

My dear friends have fueled me with courage, confidence, and love. I thank Dr. Lijun Li for keeping me company every step of the way; I thank Dr. Lixin Zhang for sprinting with me to the finish line; I thank my cohort members and co-workers for pleasant collaborations and mutual learnings; I thank Yanzhu Ji for sharing with me life wisdom and treatment suggestions; I thank Peng Li for all the delicious food and thoughtful care; I thank my dearest old friends and everyone who delivered support and comfort during my most challenging period; I thank all the sisters at ILDC and TCCDC with whom I shared the joy of dancing together.

I am extremely grateful to my family—my parents and my husband, for their love, patience, understanding, caring, and continuing support. In all these years, we have suffered from changes and difficulties that none of us would anticipate. After all the sorrow and the pain, we are still there for each other with more profound love. Now, the weight of achievement in my hands makes everything rewarding.

Abstract

There is a need to address disparities in the underrepresented minority (URM) STEM students' degree attainment and retention, vs. non-URM students. A literature review suggests that URM STEM students face barriers related to demographic, academic, and social-cognitive factors. The Louis Stokes North Star STEM Alliance (LS-NSSA) seeks to address these factors and promote URM student success in STEM disciplines. However, there is a gap in knowledge regarding LS-NSSA's mechanism and outcomes. The current study investigates factors influencing URM STEM students' retention and graduation, investigates LS-NSSA program effectiveness, and investigates the path by which outcomes are achieved. A theory-based evaluation approach is utilized in an explanatory sequential mixed-methods design guided by LS-NSSA's theory of change. The results of three sub-studies suggest the importance of first-semester experience and academic outcomes for URM STEM students' graduation and retention. Participation in LS-NSSA is associated with a higher level of academic preparation to pursue a bachelor's degree, compared to URM students not affiliated with the program. URM STEM students participating in LS-NSSA's research mentorship programs have higher levels of confidence, interest, science identity, sense of belonging, and commitment to STEM. The current study offers evidence supporting LS-NSSA's theory of change regarding student academic and social-cognitive trajectories. The current study suggests directions for future evaluative studies of LS-NSSA.

Keywords: URM STEM students, retention and degree attainment, mixed methods, theory-based evaluation, theory of change

Table of Contents

List of Tables and Figures

A Mixed-Methods Theory-based Evaluation of a Program Supporting Underrepresented Minority STEM Students

CHAPTER I: INTRODUCTION

Evaluation is critical for identifying what works and what does not work in a program. Many scholars and evaluators have interpreted evaluation as systematic inquiry (e.g., Patton, 1987; Scriven, 1991; Rossi & Freeman, 1993; Short et al., 1996), presenting it as a thoughtful, deliberate, structured, and rigorous approach (Davidson, 2004) using one or several social science research methods. The purpose of evaluation varies among scholars and practitioners in different aspects of the process. Scriven (1991) and Davidson (2004) emphasized evaluation's function of determining value or significance, including merit and worth of the evaluand (the object of an evaluation) or the product of that process. Merit refers to the absolute, internal value, while worth refers to a relative, external value (Scriven, 1991). Therefore, value judgment differentiates evaluation from pure scientific research. Chen and Rossi (1980, 1983, & 1987) argued that evaluation should achieve scientific credibility and practical worth, taking care of both effectiveness and efficiency. Pawson and Tilley (1997) related program mechanisms and results of program evaluation to the context. Following a philosophy of pragmatic liberalism, Cronbach et al. (1980) favored piecemeal inquiries that provided multifaceted perspectives. According to this view, evaluation is not directly responsive to program theory or stated goals but responds to stakeholder concerns. Weiss (1997) argued that the purpose of an evaluation is to inform the development of policy and practice, thus evaluators serve the policy and decision-making based on the evidence of the program's theory. Patton (1987) suggested that evaluations should include recommendations for program improvement, addressing the learning nature of evaluation. Theorists have emphasized that evaluation in complicated political contexts is never value-free but should serve the aims of social justice, human rights, empowerment, and equity (Greene, 2006; Howe, 2003; Mertens & Wilson, 2018; MacDonald, 1974). However, these values and goals cannot be realized without the effective use of social resources aligned with an appropriate purpose.

Evaluation is complex, and its process and purpose remain distinctive. In the current study, focused on STEM education, evaluation is understood as a multifaceted systematic inquiry into an intervention, including its theory of change, activities, outputs and outcomes, context, and value. Evaluation enables integrated, evidence-based information through the use of qualitative and quantitative methods, tests its theory of change, and examines the effectiveness of an intervention. Equally importantly, evaluation is performed to determine the value of the evaluands, inform and support decision-making, understand the impact on stakeholders, and promote social justice. Such as such, evaluation seeks ways to translate the evaluation results into practice.

In STEM education, evaluation employs rigorous measures to obtain evidence to inform educational policies, practices, and interventions. Evaluation is a vital managerial and learning tool. instrumentally used to demonstrate the extent to which an intervention achieves its goals in responding to students' needs; improving students' academic performances; institutionalizing existing projects to serve more students; shaping the possible strategic directions and approaches to balance the allocation of educational resources in STEM disciplines; informing the application of the intervention to other populations in other contexts. By involving diverse stakeholder groups into evaluation activities, evaluation allows them to define and address issues standing on their grounds. This is especially important for historically marginalized populations and underrepresented minorities. Therefore, the current study shares the experience of how an internal evaluator can apply a mixed-methods theory-based evaluation approach to enrich the understanding of a program supporting underrepresented minority STEM Students.

Problem Statement

The topic of the persistence, retention, and graduation of minority students in the science, technology, engineering, and mathematics (STEM) fields in the United States has been widely discussed, in part owing to the call to increase the size of the STEM workforce (National Science Foundation, 2011; National Science Board, 2018). The U.S. Department of Labor has predicted a gap of approximately 1.6 million workers with degrees in computing sciences between the available employee pool and national demand (Charleston et al., 2014). The 2012 report of the President's Council of Advisors on Science and Technology (PCAST) emphasized the need for a STEM-educated workforce prepared to enhance the country's competitiveness in the global marketplace (Olson & Riordan, 2012). National efforts have been launched to diversify those working in scientific enterprises, such as the Innovate Initiative of the America COMPETES Act of 2007 and 2010, and these have all targeted increasing participation in STEM careers (Byars-Winston et al., 2016). This topic is critical. Over a long period, the number of college students who enroll in STEM majors has dropped by 50%, and approximately half of those who do enter STEM programs transfer out of their STEM major or their postsecondary institution before completing their degrees (Simon et al., 2015). The graduation rates of students in STEM majors remain roughly 20% lower than that of their counterparts in non-STEM majors (DeAngelo, et al., 2011; Chen, 2013).

This is not simply an educational issue, but also an issue related to social justice and racial inequality. The disparities between underrepresented-minority (URM) and White or Asian-American students in STEM are even more serious. Although the number of STEM jobs have grown faster than non-STEM jobs since 2010, and many STEM jobs are forecasted to grow continuously in the future decades, this growth has been unevenly distributed across racial groups. Underrepresented minority population collectively represented 30% of the employed population in the United States but 23% of the total STEM workforce (National Science Board, 2022). URM students in STEM attain degrees at a much lower rate than non-URM students (NSB, 2018), which is consistent with the fact that even though URMs have a lower likelihood of enrolling in college, they experience a higher rate of attrition from postsecondary institutions (Xu & Webber, 2016; Whitcomb & Singh, 2021). For decades, STEM careers have been described as "nontraditional" for women and culturally underrepresented groups, reflecting the fact that such individuals show reduced pursuit of relevant occupations, due to perceived barriers (Milner et al., 2014). There is, therefore, an extensive mismatch between the cultural groups, as found by Byars-Winston, et al. (2016), and it persists despite ongoing efforts to increase the participation of URMs in the STEM fields. This argument is supported by the statistical evidence: though more than half of freshman URM students report that they intend to major in a STEM field (Olson & Riordan, 2012), together they earn only about 17% of the bachelor's degrees granted in those fields (National Center for Education Statistics, 2018). Because the non-White population is expected to increase, making up over 50% of the U.S. population by 2050, their success in STEM fields is inextricably connected to national scientific innovation and economic prosperity (Xu, 2018).

Statement of the Context

In December 2018, the Committee on STEM Education (CoSTEM) of the National Science and Technology Council (NSTC) published the federal government's five-year strategic plan for STEM education, calling for a nationwide collaboration with learners, families, educators, communities, and employers for the STEM community (National Science and Technology Council, 2018). Specifically, one of the plan's three central goals is increasing diversity, equity, and inclusion in STEM, which is key for achieving the other two, because diversity in the workforce provides an inclusive environment that better retains talent and promotes productivity (NSTC, 2018). In October 2019, the Office of Science and Technology Policy (OSTP) issued an annual progress report regarding efforts related to this strategic plan. In total, 16 federal agencies enacted 125 programs, with an estimated budget of 3,203 million U.S. dollars in fiscal year 2019 (OSTP, 2019). Among these agencies, more than one-third of investments were funded by the National Science Foundation (NSF), which operated 24 programs in 2019.

NSF LSAMP Program

NSF has a long history of supporting students' development and success in STEM fields. In 1991, NSF initiated the Louis Stokes Alliances for Minority Participation Program (LSAMP), aimed at increasing the quality and quantity of students successfully completing STEM baccalaureate degree programs (NSF, n.d.). LSAMP advocates innovative, evidenced-based recruitment and retention strategies to support racial and ethnic groups historically underrepresented in STEM disciplines. Grant recipients are projects conducted through alliances of academic and relevant institutions. These projects are expected to implement sustained and comprehensive approaches that facilitate long-term academic and career achievements among URM STEM students. LSAMP supports activities related to critical transition points in STEM education, including secondary to postsecondary study, two-year to four-year institutions, undergraduate to graduate study, and undergraduate to career at workplaces (NSF, n.d.).

LS-NSSA: The Target Program

The North Star STEM Alliance (LS-NSSA) is a member of LSAMP. Located in the University of Minnesota (UMN), this Alliance currently includes four two-year colleges, ten four-year universities and campuses (including public colleges and universities, private colleges, and tribal colleges), and three community organizations. Two institutions will be added in the next grant cycle. This statewide partnership is committed to supporting URM STEM students. NSF has renewed the five-year funding twice since 2007. The latest grant covers program operations from July 2017 to July 2023, including a one-year extension.

The Alliance partnership has been growing numerically and geographically. Partners establish cohort and study programs, fund undergraduate research, and support students' attendance at local and national STEM conferences. The signature practices of this program include undergraduate research projects, cohorts and community-building activities, faculty and peer mentoring programs, academic and financial supports, and professional and career development workshops and seminars. To include the voices from different STEM disciplines and the STEM community, this program engages coprincipal investigators and faculty members from partnering institutions' STEM colleges, the UMN Office of Undergraduate Education, the UMN Office for Equity and Diversity, and the UMN Office of Academic and Student Affairs.

Although retention and graduation in STEM majors has increased in the past ten years, there is urgency in the Alliance to improve the experience and retention in STEM disciplines. The Alliance's partnership is an important convener for exploring and addressing the complex set of factors that affect URM STEM students' retention and graduation, informing the effectiveness of the program, and understanding and improving the experience of URM STEM students. The steering committee of the program holds annual site coordinator meetings, transfer program stakeholder meetings, and governing board meetings to strengthen the partnership and delve more deeply into concerns about URM STEM students' academic success.

Program Evaluation Efforts

Purpose of Evaluation. As the Alliance has been funded by LSAMP for 16 years (three five-year phases plus one year of extension), the program is now recognized by NSF as a well-established program grounded in sound programmatic approaches (see Appendix A. The Phase 3 Logic Model). For program improvement and evidence generation, LS-NSSA's program evaluation consists of two parts—the internal evaluation and the external evaluation, both responding to program goals and objectives (see Appendix B. The Phase 3 Evaluation Plan, 2017-2022). An external evaluation has been conducted annually. Evaluation specialists from the Center for Applied Research and Educational Improvement (CAREI), University of Minnesota conduct annual surveys for site coordinators, governing board members, and community partners to investigate opinions regarding program implementation, partnership development and collaboration, program outcomes, program progress, institutionalization, and sustainability.

The purpose of the internal evaluation is to empirically describe the progress and outcomes of the program, explore the patterns of URM STEM students' engagement and persistence, and test participants' achievements. The internal evaluation addresses the value of the program from student participants' perspective and explores students' experiences. The internal evaluation tracks annual benchmarks to measure progress in the number of URM STEM graduates, including enrollment data and degree data across the Alliance and distinctive interventions developed by partner institutions, and explores factors influencing retention and graduation. To assess changes related to diversity and inclusion on campus, the internal evaluation measurements focus on the engagement of URM STEM students in co-curricular programs and students' perceived inclusion and barriers.

Existing Evaluation Efforts. Between 2008 and 2021, the Alliance directly supported (i.e., provided financially support in any way) 4,682 URM undergraduate students majoring in STEM. To serve these students through multiple opportunities, the Alliance engaged about 200 faculties and staff members annually and connected almost 800 URM students with undergraduate research opportunities. Over 1,400 students attended local, regional, and national academic conferences through LS-NSSA's financial support (LS-NSSA, 2020).

Enrollment and degree attainment trends have been tracked since the beginning of this program. The WebAMP data reporting system is used by all partner institutions to upload institutional data. Enrollment and degree attainment trends are depicted in 11 clusters of STEM disciplines (e.g., Agricultural Science, Biological Sciences, Computer and Information Sciences, Engineering, Mathematics, Physical Sciences, etc.) over time, by gender, race/ethnicity, and STEM fields. The results showed that the enrollment of URM STEM students is on an upward trend, from 907 in 2008 to 2,985 in 2021, with a projection of 3,009 in the 2022 academic year. The gains occurred across nearly all cultural groups including Native Americans—the least engaged group in previous grant periods (Phase 1 and Phase 2). STEM degree attainment also continues to increase. In the first five years of the partnership, the Alliance exceeded the goal of doubling the number of URM STEM graduates from the baseline of 136 in the academic year 2004–2005 and welcomed 326 graduates in the seventh year. In the 2021 academic year, the number reached 675 with a projection of 733 in the 2022 academic year. The gains occurred across nearly all URM ethnic groups.

An internal evaluator examined demographic factors related to the graduation of URM freshmen who enter with a STEM academic plan (Soria, 2014). The models incorporated eligibility for Pell Grants, first-generation status, race/ethnicity, and ACT scores. ACT scores were positively associated with URM STEM students' four- and five-year graduation rates ($p < .001$ and $p < .05$, respectively). Besides, URM STEM students enrolled from high schools were more likely to extend one to two years than their non-STEM counterparts ($p < .05$).

Additional insights into the program's effectiveness were derived via qualitative research. Together with quantitative results, this knowledge informed the development of the logic model. In Phase 1 and Phase 2, LS-NSSA's internal evaluators interviewed program participants to explore the components of the program that were most useful to participants; students' experiences related to campus climate; students' experiences in undergraduate research; students' interactions with faculty, peers, and graduate students; students' post-graduation career aspirations; measures campus administrators could take to increase students' sense of belongingness; and barriers to students' participation in this program. The results affirmed the value of the program, highlighted the need for a higher level of support for first-generation student and students who are eligible for the Federal Pell Grant, and the need for strengthening faculty and peer mentorships (Soria, 2014; Mixson, 2015; Hornickel, 2016). Regarding undergraduate research activities, some evaluation and social science studies led by internal evaluators and UMN faculty members clearly showed that many participants have been engaged in a succession of research experiences through LS-NSSA, and community college students' engagement in research increased markedly (Cain, 2017). Despite challenges and barriers experienced by URM STEM students, LS-NSSA's research resources and connections mattered for students' success (Jehangir et al., 2022).

The Current Study

Existing evaluation studies and results guided development of a suitable logic model. Based on this knowledge, the LS-NSSA program chose to focus on URM college students enrolled in STEM disciplines, especially first-generation students and those who need financial support. This focus sought to ensure efficient resource allocation. However, questions remained about how, and to what extent, the LS-NSSA program worked. There was a gap in knowledge regarding program elements that sought to promote successful conversion from program activities to the increased URM STEM degree attainment.

The current study reports results from internal evaluation studies conducted by the author that seek to fill these gaps in knowledge. The current study consists of three sub-studies. To clarify and enrich the program's theory of change, the current study starts with an inquiry into the theory base of the intervention, probing factors that influence the target population's retention and graduation (sub-study 1). Then, it tests whether and how program participants outperform their non-participant counterparts (sub-study 2), followed by an exploration of students' research mentoring experience and their demands regarding diversity, equity, and inclusion (sub-study 3). Together, these studies seek to establish a better understanding of LS-NSSA's effectiveness and inform ongoing program implementation and development.

In addition, the current study seeks to address equity in STEM education utilizing an evaluative lens. It applies a theory-based evaluation approach, focusing on LS-NSSA's theory of change, the intervention's mechanisms, the values perceived by participants, and understanding of the program's specific context. It expands the use of the theory-based evaluation approach in a formative internal evaluation by engaging the lens of critical race theory and utilizing mixed-methods.

Research Questions

The current study addressed the following research questions:

1) What are the factors that influence the retention and degree attainment of LS-NSSA's target population (URM STEM undergraduate students)?

2) How do LS-NSSA participants perform on key factors compared with their non-participant counterparts?

3) How do participants perceive the influence of LS-NSSA, specifically, the research mentorship programs?

a) How does the experience at LS-NSSA influence their confidence, interest, belongingness, and commitment in STEM fields?
b) How do participants perceive LS-NSSA's influence on diversity, equity, and inclusion?

The research questions and the design of the study are based on the hypothesis that (a) undergraduate URM students' success in the STEM fields is associated with multiple factors; (b) understanding of program mechanisms and contextual factors may inform future program design to support URM STEM students; (c) URM STEM students encounter barriers and difficulties generated by their compound identity and experience, which may be different from the identity and experience of non-URM STEM counterparts; (d) undergraduate research and faculty mentorship play important roles in the target population's learning experience and retention; (e) experiential knowledge of participants is as important as quantitative metrics in exploring and evaluating the perspectives and program experiences of culturally diverse individuals; (f) a theory-based evaluation approach through a critical race theory lens is appropriate for this task.

Definition of Terms

Alliance – The consortia of 14 participating educational institutions and three community organizations.

DEI – Diversity, equity, and inclusion.

Evaluand – The generic term for the object that is being evaluated, including person, performance, product, project, program, intervention, policy, and so on (Scriven, 1991).

LSAMP – The Louis Stokes Alliance for Minority Participation (LSAMP) program, initiated by the National Science Foundation (NSF) in 1991, which aims at increasing the quality and quantity of students successfully completing STEM degrees.

LS-NSSA – The Louis Stokes North Star STEM Alliance, a member of LSAMP in the State of Minnesota.

MnDRIVE – Minnesota's Discovery, Research, and InnoVation Economy program, a partnership between the University of Minnesota and the State of Minnesota that funds research and studies addressing emerging challenges in STEM fields.

NSF – National Science Foundation.

Persistence – Students return to college at any institution for the following academic year.

Retention – Students return to college and register at the same institution for the following academic year.

STEM students – Students who declare a STEM major when they arrive at the university. STEM majors are defined by the NSF STEM Classification of Instructional Programs Crosswalk (NSF, 2018), consistent with the STEM Designated Degree Program list generated from the Department of Education's Classification of Instructional Programs (CIP).

URM – NSF defines the underrepresented minority (URM) population to include Hispanic/Latino, Black/African American, American Indian and Alaska Native, Native

Hawaiian and Other Pacific Islanders, and those of two or more races with at least one being underrepresented. Notably, Southeast Asian students are excluded from this definition.[1]

Virtual Ep. – The Victual Epidemiological Research Program that provides spring and summer research opportunities for LS-NSSA's target population.

Chapter Summary

URM STEM students' retention and graduation are critical issues in STEM education. The disparities between URM and White students in STEM are serious. There is a need to address and solve this issue in the higher education system. LS-NSSA is a 15-year-old program that supports URM STEM students by connecting students with financial resources, mentorships, and professional development opportunities. LS-NSSA seeks to respond to the need to promote URM student success in STEM disciplines. Phase 1 and Phase 2 evaluation results have provided a broad picture of LS-NSSA's importance, progress, outputs, and outcomes. However, there is a gap in knowledge regarding elements that mediate the relationship between program activities and program outcomes. There is a need for more information. The current study addresses this need by focusing on three research questions regarding the program's effectiveness and value, utilizing a theory-based evaluation approach and mixed methods.

[1] The Alliance utilizes alternative sources of grant funding to support this population.

CHAPTER II: LITERATURE REVIEW

The literature review focuses on the background and issues that shape the target program and the current study. The four major topics addressed in the current study are 1) factors that influence URM STEM students' retention and success; 2) programming for enhancing URM STEM students' outcomes; 3) application of theory-based program evaluation to URM STEM student programming; 4) application of critical race theory to the topic of URM STEM student retention and success. This review summarizes existing theory and research findings regarding each topic.

The review of diverse factors that influence the retention and graduation of students in STEM disciplines informs the exploration of factors that influence URM students in the Alliance context. The exploration of existing programming and interventions that support URM STEM students provides context for LS-NSSA's work. The review of literature regarding the theory-based evaluation approach frames the rationale for utilizing this approach in the evaluation of LS-NSSA. The review of critical race theory and its application in higher education guides the inquiry regarding LS-NSSA student participants' program experience. Together, the literature review informs the utilization of the theory-based evaluation approach to identify key metrics for evaluating LS-NSSA within a culturally diverse context. The current study seeks to build knowledge regarding effective interventions for URM STEM students.

Factors Influencing the Retention and Graduation of Students in STEM Fields

Recent findings indicate that STEM students' retention and graduation may be affected by a variety of factors including: (a) Challenges related to demographic characteristics such as gender, race/ethnicity, socioeconomic status, and family education history; (b) academic performance; (c) social-cognitive factors (e.g., participation, engagement, identity development, self-efficacy, and belongingness).

Challenges Related to Demographic Characteristics

Quantitative and qualitative studies have explored the relationship between demographic characteristics and students' outcomes in STEM fields. Concern for paying for college limits the social and academic adjustment of URM STEM students. Socioeconomic status is a significant concern for STEM students, especially for those who transfer from community colleges to four-year institutions (Xu, 2018). Anderson and Kim (2006) pointed out that at the three-year point, all ethnic groups were almost equally likely to continue STEM enrollment, but following the third year, the progress of African Americans and Hispanic/Latino students slowed compared with their counterparts. They are more likely to work more hours on- and off-campus during college than their counterparts, which may inhibit retention (Hurtado, et al., 2010).

Apart from financial challenges, many URM STEM students recognize that characteristics such as gender, race, and ethnic background may be associated with adverse experiences that depress graduation rates. Some positive effects are even mediated by gender and race factors. For example, although it has been widely known that students who engage more frequently in on-campus student services are significantly more likely to have aspirational momentum, as Belser et al. (2017) addressed in their study, this relationship is weaker among female students and Black or Latinx students, indicating that these groups face more challenges in their STEM learning experience (Jackson, 2013; Charleston et al., 2014; Morganson et al., 2015; Wang et al., 2017). Wang et al. (2017) found that being female and being Black is positively associated with a student's intention to transfer to a non-STEM major after being enrolled as a first-year STEM student. Riegle-Crumb et al.'s (2019) study confirms that URM STEM students experience a higher frequency of setbacks in pursuing a bachelor's degree, and the STEM field has disproportionately excluded URM students. Across undergraduate disciplines, STEM is the only field where Black and Latinx students are significantly more likely than their White peers to either switch directions from their original majors to a non-STEM field or drop out from their institutions. This is partly because URM students recognized that there were damaging misperceptions, stereotypes, and microaggressions about their academic and intellectual abilities as a result of their identity (Charleston et al., 2014). In Allen et al.'s (2022) study, Black female participants described numerous experiences with sexism and racism in creating a hostile environment that discouraged their passion in STEM.

Racial microaggressions—subtle acts of racism and explicit or implicit messages—are ingrained in the campus and classroom climate, resulting in social, emotional, and health stresses for students of diverse cultural backgrounds (Sue et al., 2008; Lee et al., 2020; Robinson-Perez et al., 2020; Smith et al., 2022). Lee et al. (2020) found that Black students in STEM majors are more likely to experience racial microaggressions when interacting with instructors, advisers, and peers than other students of other cultural groups. For Latinx students, both male and female students experienced microaggressions, especially stereotypes in their intelligence and academic success (Smith et al., 2022). Robinson-Perez et al. (2020) found that the level of psychological distress such as anxiety and depressive disorders tends to be higher among students of diverse cultural backgrounds who were experiencing a low-achieving/undesirable culture. This finding echoes Torres-Harding et al.'s (2012) argument that students from culturally diverse groups were made to be self-doubted, believing that they were incompetent and low achieving, and any of their achievements must result from unfair entitlements and special treatment. Psychosocial threats generated from the negative campus and classroom climate contribute to the remaining achievement gap between URM students and their White counterparts (Jordt et al., 2017).

Of all students in the United States who finished a four-year degree in the academic year 2010–2011, 45% had previously enrolled at a two-year institution (Jackson et al., 2013) and 74% had attended a community college to complete credits towards a bachelor's degree (Tupper et al., 2010). Community colleges serve as the entry point for many URM students to reach postsecondary education (Jackson, 2013). Transferring from community colleges to four-year institutions should have been a promising pathway for students who want to pursue a bachelor's degree in STEM. However, this pathway makes no easy steps for URM students. In general, women and students with diverse cultural backgrounds persist in and transfer into STEM fields at rates that remain lower than those of their White male counterparts (Wang, 2016). Wang et al. (2017) provided an explanation for this disparity: many women and students of diverse cultural backgrounds at community colleges have obligations to their families and full-time jobs in addition to their academic pursuits, and these are major barriers for them, as are financial concerns. Reyes (2011) and D'Amico et al. (2014) listed certain unique challenges for URM transfer students in STEM regarding gender and cultural background, including their status as head of household, childcare responsibilities, parental educational level, and perceived discrimination based on the intersections of gender, ethnicity, and age. As could be expected from the results noted previously, being female and being Black/African American are associated with a greater transfer shock—the difference in transfer students' cumulative grade point average (GPA) from the most recent transfer institution and first-semester GPA at the focal institution (Lakin & Elliott, 2016; Reyes, 2011; Allen et al., 2022). There are three additional demographic characteristics that are negatively associated with URM STEM transfer students' success: 1) being a STEM transfer student of a non-traditional age—younger students succeed as transfer STEM students at higher rates; 2) being a first-generation college student—the outcomes of first-generation URM students are also more influenced by the college environment before and after transfer; and 3) speaking English as a second language (Wang, 2016; Crisp & Nuñez, 2014; Dika & D'Amico, 2016; Wang et al., 2017).

Academic Factors

Academic performance and achievements reflect students' learning conditions. In the case of students starting their STEM academics at four-year institutions, Belser et al. (2017) provided numerical evidence that supports the importance of major declaration: students initially classified as "STEM majors" at entry are 15 times and 18 times more likely, respectively, to retain their STEM major in their second year of college, compared to students initially classified as "major-undecided." An early career commitment and the effort in developing career readiness also matter. students who participated and completed requirements in the career planning course were three times more likely to be retained in the second year of college, and every one-unit decrease in the Career Thoughts Inventory (CTI) score (i.e., career unreadiness measurement) doubled the odds of being retained in a STEM major (Belser et al., 2017).

For STEM students of all demographic groups, previous research has demonstrated the importance of the first semester and first-year GPA on their persistence within six years

(Rogulkin, 2011). Beyond the first year, when students' cumulative GPA is lower than 2.5 later in college, the probability of dropping out of the institution becomes higher compared with students with a GPA of 3.5 or higher (Chen, 2013). Pedraza and Chen (2022) examined the relationship between academic predictors and degree attainment. The result shows that among various predictors, the highest correlation occurs between students' overall GPA and the degree outcome with a moderate effect size of 0.46, followed by their first-year GPA (effect size = 0.38) and internship experience (effect size = 0.34). Other predictors display weak correlations with degree attainment, including experience in undergraduate research, studying abroad, service-learning, and mentorship involvement.

Through a case study in a Hispanic-Serving Institution (HSI), Ortiz and Sriraman (2015) pointed out that faculty members' top-selected academic factors that influenced a URM student's retention and persistence in STEM disciplines were: student's academic achievement (e.g., GPA) and their quantitative skills (i.e., computer technic skills and math ability). This argument is supported by Simon et al.'s (2015) study, using structural equation modeling. They suggested that a higher level of academic achievement is associated with a greater likelihood of persistence in the Science domain, mediated by students' self-autonomy and emotional well-being. Regarding academic achievement, Dika and D'Amico (2016) had a deeper investigation and suggested that a better first-semester GPA increases the odds of first-generation URM students' three-year persistence across STEM majors. They also indicated that students' preparation in math is a significant factor in their persistence in Physics, Engineering, Mathematics, and Computer Science (PEMC) majors. On the contrary, a lower level of student's cumulative GPA, together with the negatively perceived academic quality and less support from faculty members, is significantly related to students' intention to drop out of the institution. However, a lower cumulative GPA is not associated with the major-changing decision by itself if the quality of the academic program and the accessibility of faculty support are guaranteed (Xu, 2018). This finding indicates that students' learning experience is as important as and predicted to academic achievements to some extent. For example, students who are struggling in STEM majors need to meet with academic advisors and participate in study groups more frequently (Gayles & Ampaw, 2014); smaller class sizes and stronger relationships with faculty members are considered significant to Black and Latinx students' STEM achievement (Green et al., 2019); engaging students in active learning activities (i.e., classroom activities that are meaningful to students' course or career planning) is significantly associated with students' confidence in persisting in STEM education (Wang et al., 2017); URM students can benefit from involvement in STEM-specific academic and professional organizations that enhance their academic integration (Chang et al., 2014; Espinosa, 2011).

Quantitative skills such as basic mathematics and science competency are fundamental skills required in STEM disciplines. Eris et al.'s (2007) study on Engineering students found that students who leave Engineering within two years of enrollment (usually women and URM students) have lower levels of competency in their math and science skills than students who persist. Lack of preparation for collegiate mathematics and science among URM STEM students and the difficulties that they experience in collegiate mathematics courses have undermined their enrollment and success (McGee & Martin, 2011).

Regarding URM STEM transfer students, the work of D'Amico et al. (2014) suggests that there are a variety of academic factors that predict early integration and academic outcomes in transfer students' first-year grades and retention in the following semesters. Significant indicators include the attainment of an associate degree, early major declaration, type of the original institution, students' preparedness in math and writing, time since last college enrollment, and GPA prior to transferring. While transfer GPA in itself is negatively associated with perceived early academic fit, and every one-point increase in GPA at the community college is only associated with a 0.71-point increase in academic performance in the four-year institution (Lakin & Elliott, 2016), demonstrating a higher transfer GPA does not guarantee complete adjustment to the setting of the new university. First-semester GPA at a four-year institution is a vital indicator predicting the persistence of STEM transfer students (Dika & D'Amico, 2016). Lakin and Elliott (2016) observed an average decline in GPA after a transfer of 0.63 points. In general, a 1-point decline in GPA leads to a four percent greater likelihood of leaving the college of enrollment and a doubled likelihood of leaving the institution. Tupper et al.'s (2010) observations of the Grace Hopper Scholar Program (GHSP) also indicate that the dropout rate of part-time students is almost twice that of students enrolled full-time or who have mixed status.

Declaring a STEM major earlier keeps transfer students in STEM fields (Xu, 2018; Belser et al., 2017; Crisp & Nuñez, 2014). By contrast, taking more credit hours in the first semester after transfer and taking more STEM credits (more than 40%) slightly decreases the likelihood of a change in academic major or leaving the institution (Lakin & Elliott, 2016; Dika & D'Amico, 2016; Wang et al., 2017; Xu, 2018; Ortiz & Sriraman, 2015).

Wang's (2016) research finds that course-taking patterns in math courses during the first term in community colleges are not the most frequent factor determining success among STEM transfer students. Otherwise, taking likely transferable STEM courses is the most viable trajectory for STEM transfer. To accomplish a successful vertical transfer, students should successfully complete at least three credits of likely transferable STEM courses and complete as many STEM credits as possible to increase the likelihood of success. Research shows that women students would only need a minimal amount of math credits to enable success in STEM transfer (Wang, 2016).

Social-Cognitive Factors

Research has long suggested that the major causes of first-year attrition are social and emotional rather than academic (Miller & Servaty-Seib, 2016), and social-cognitive factors affecting students' behaviors and choices are extensively discussed in the studies investigating educational psychology. Persistence, retention, and degree attainment are thus associated with cyclical interactions among the educational experiences, psychological perceptions, and sociological outcomes of students (Hanauer et al., 2016). A number of theory-informed measures have been implemented to assess this interaction (Chemers et al., 2011; Estrada et al., 2011; Milner et al., 2014).

Tinto's (1975) model of students' departure, later refined by Tinto in 1997, brought students' interactions with faculty members, peer-group relationships, involvement in extracurricular activities, and goal commitment into focus. Students' integration into the formal and informal academic and social life in college were engaged as important social-cognitive factors regarding their retention and degree attainment. Nowadays, many interventions and programs for preventing students' early dropout in universities and colleges still ground their designs in this model (e.g., Carpi, Ronan, Falconer, Boyd et al., 2013; Sweeder et al., 2021). In 1990s, social cognitive career theory (SCCT; Lent et al., 1994) was developed based on Bandura and National Inst of Mental Health's (1986) general social cognitive theory. It connects social-cognitive factors with career development and clarifies the psychological process in students' behavioral choices, making it one of the fundamental theories used to address the challenges and retention of students in their majors and institutions, including STEM students. Self-efficacy, science identity, belongingness, outcome expectations, goal commitments, values, and interests have been the main concepts at play here, and they are interrelated with the environmental factors, institutional practices, academic factors, and students' demographic characteristics.

Self-efficacy and learning experiences are central constructs in exploring how basic academic and career interests are developed, educational and career choices are made, and academic and career success is obtained (Lent et al., 1994). In detail, four main aspects contribute to self-efficacy: performance accomplishments (e.g., personal mastery experiences or past successes), vicarious learning (e.g., observing the explicit behaviors of role models), social persuasion (e.g., verbal encouragement), and effective or emotional arousal experienced while completing a task (e.g., low anxiety and relaxation) (Byars-Winston et al., 2016). Studies conducted by Hurtado et al. (2010), Chemers et al. (2011), Estrada et al. (2011), Byars-Winston et al. (2015), Hanauer et al. (2016), and Syed et al., (2019) focuses on the contributions of self-efficacy to the outcomes of STEM majors among students of diverse cultural backgrounds. The results show that scientific and research-related self-efficacy—confidence in their own ability to successfully perform scientific work or conduct research—is the mediator between students' research and learning experiences and their commitment to a research career. Syed et al. (2019) also differentiated the science self-efficacy and the leadership/teamwork self-efficacy and demonstrated the combined mediation effects to the STEM career commitment. Some STEM students cite their strong supporting networks as helpful for surviving and thriving. However, statistical analysis has found the opposite effect of a supportive network for students in community colleges. Receiving emotional support from one's family and friends appears to have a negative effect on students' intention to transfer into bachelor's degree programs in STEM fields, partly explained by a greater sense of belonging in the community colleges (Wang et al., 2017). Being a first-generation college student is also negatively associated with transfer self-efficacy (Wang et al., 2017).

Having a scientific identity strengthens the confidence of STEM transfer students in the development of their STEM careers. Higgins et al. (2011) suggested that students in science majors should be recognized and treated as scientists instead of technicians. Where students are able to view themselves as members of the STEM enterprise, they can commit to challenges and overcome the obstacles resulting from their identification within the field (Jackson et al., 2013). URM STEM students are more aware of negative psychological implications. Byars-Winston et al. (2016) presented a framework of scientific identity, consisting of three overlapping dimensions, including competence, performance, and recognition. Belongingness is the representation of such identity, and a sense of isolation reflects the lack of it. Class belongingness contributes to the engagement in STEM coursework most consistently, which is related to the classroom environment including connection to instructors and peers (Allen et al., 2022). However, a sense of isolation and invisibility, along with the academic pressures of larger classes, often lead women and

URMs to switch out of STEM majors (Jackson et al., 2013). Their isolation within the academic environment bears a resemblance to the isolation they experience in daily life (Charleston et al., 2014; Allen et al., 2022). Yosso et al. (2009) found that institutional microaggressions and the lack of racial/ethnic diversity of the faculty have led to feelings of isolation and hopeless across Latinx students. For URM transfer students, comfort and familiarity with the previous community college slightly limit academic and social connections within the first few weeks at the receiving university (D'Amico et al., 2014). Interestingly, the perceived social fit is not a significant positive predictor of GPA for them than for non-transfer URM STEM students, which supports the argument that transfer students' adjustment on campus is more associated with academics than with the social reception.

Researchers have also employed theories other than SCCT. The motivational model of Simon et al. (2015) and Perez et al.'s (2014) cost-benefit analysis explores personal values in-depth as well as the trade-off between perceived costs and benefits, which drives the development of self-determination, in contradiction of the assumption that students' choices and behaviors are mandated or solicited by their social environment. Perez et al. (2014) concluded that "higher engagement in exploring options before committing to a career path was likely to start the academic journey with the feeling of competent, perceiving higher value, and believing that investing effort in the STEM major was worthwhile" (p. 324). Perceived costs play a primary role in students' leaving intentions, but students weigh the values of their majors against the costs after gaining experience and receiving feedback in courses. Morganson et al. (2015) used embeddedness theory (Mitchell et al., 2001) to engage a broader community context. They find that solidarity and friendship with peers in STEM majors, role models, and cultural icons are positive factors for building up a sense of community embeddedness: passion, challenge, skills, aptitude, concrete tasks, real-world applications, and intrinsic outcome expectations are positive for major embeddedness. Critical discourse analysis allows for an in-depth examination of the language used as a type of social practice that both reflects and constructs the social world and affects an individual's embeddedness process and identity development. It has revealed that students of diverse cultural backgrounds find more difficulty in being embedded where they are construed as problems—either as being underprepared or at risk (Castro, 2014).

Interventions Enhancing URM STEM Students' Outcomes

The resources and activities provided by institutions play a critical role in maintaining and boosting students' success in STEM education. Many programs in universities are supported by the National Science Foundation (e.g., Tupper et al., 2010; Palmer et al., 2011; Reyes, 2011; Higgins et al., 2011; Olson & Labov, 2012; Kendricks et al., 2013; Drew et al., 2016; Cott et al., 2016), and they employ similar evidence-informed approaches that are nevertheless distinct. Among these, financial support, course-based undergraduate research, peer mentoring, and faculty mentoring have been most widely implemented and proven to have a strong positive influence on URM STEM students (e.g., Reyes, 2011; Higgins et al., 2011; Olson & Labov, 2012; Kendricks et al., 2013; Drew et al., 2016; Cott et al., 2016; Carpi et al., 2017; Syed et al., 2019).

Financial Support for URM STEM Students

To overcome barriers related to financial burdens, students rely heavily on technical resources and other institutional structures to facilitate an effective learning process. A quasi-experimental research project identifies that, in general, eligibility for need-based financial aid is positively associated with STEM credit completion by 20% to 35% among STEM students (Castleman et al., 2018). In addition to federal financial aid such as Pell Grant, in recent decades, many four-year institutions and community colleges have provided direct financial aid or conditional financial support for alleviating URM STEM students' pressure from outside sources (Wright et al., 2021), encouraging them to stay on track, improving their commitment and science identity development (Oseguera, et al., 2020), and enhancing their academic performance and retention (Jackson et al., 2013; Wang, at al., 2017).

For example, in 1999, Louisiana State University (LSU) initiated the Computer Science, Engineering, and Mathematics Scholarships (CSEMS) program—a pure financial aid program for high-quality students struggling with finance. Responding to NSF's changing grant requirement, CSEMS evolved into the Scholarships for Science, Technology, Engineering, and Mathematics (S-STEM) program in 2004 by integrating additional professional development services, advising strategies, and mentorships. Research shows that 94% of S-STEM program participants retained in STEM fields, much higher than non-participant low-income URM STEM students (34% on average; Wilson, Iyengar, et al., 2012). The University of Maryland Baltimore County (UMBC) has sought to bridge Black undergraduate students with STEM doctoral programs through its Meyerhoff Scholarship Program (MSP). MSP provides integrated services for the target population, while in a survey for the first 15 MSP cohorts, respondents rated financial scholarship as the most helpful support (Maton et al. 2009; Stolle-McAllister & Carrillo, 2011). It finds out that MSP program participants are twice as likely to obtain a bachelor's degree in STEM and five times more likely to pursue a doctoral program in STEM than their non-participant counterparts (Maton et al. 2000).

Some programs are conditional. The SPRING program based in Lyman Briggs College at Michigan State University provides student participants with 3,000 dollars, 6,000 dollars, and 9,000 dollars in their first year, second year, and third year in the program, respectively, by completing required STEM preparation courses. Sweeder et al. (2021) compared the degree attainment within six years of enrollment between participants and non-participants. The result suggests that for Pell-eligible students, participants' six-year graduation rate was 11% higher than non-participants (92% vs. 81%); for non-Pell-eligible students, the gap increased to 23% (95% vs. 72%). The NanoSTEM program at Binghamton University, serving economically disadvantaged students and transfer students in the sciences and engineering field, attaches more requirements on financial support. By fulfilling conditions such as participating in the orientation program, taking the STEM seminar class, achieving a 3.0 GPA, involving in research projects, and obtaining an undergraduate teaching assistantship, students can receive up to 8,000 dollars per academic year and summer tuition benefits (Cott at al., 2016). These requirements push students to realize their academic goals and improve their on-campus engagements.

A few studies have examined interventions within community colleges to inspire transfer. To stimulate students' motivation to take STEM courses, the Grace Hopper Scholar Program (GHSP), established at the Community College of Baltimore County in Maryland, provided each qualified student with a cash reimbursement for the completion of their first mathematics or computer science course for credit in which they received a grade of C or better, which helped 30 of 74 (41%) of GHSP participants to successfully transfer to four-year institutions (Tupper et al., 2010). In addition, Crisp and Nuñez (2014) found a positive relationship between receiving financial aid and transfer odds.

Because community colleges have only a limited capacity to offer a broad range of high-quality STEM courses and supportive technical programs (Wang et al., 2017), cooperation with four-year institutions is important for students' development. Programs supported by LSAMP grant funding have shown examples of collaboration. Program alignment between two-year and four-year institutions is key to successful transfer and students' retention in STEM majors, especially for members of underrepresented ethnic groups (Jackson et al., 2013). This alignment should be built upon consistent information regarding academic requirements or transfer requirements from varying institutional types. As Jackson (2013) illustrated, informational asymmetry generates anxiety stemming from the transfer out of one institutional culture and into a very different and complex institutional culture.

Undergraduate Research Opportunities

A considerable number of existing studies have demonstrated that engaging in undergraduate research experiences significantly increases students' understanding of science and research, skill and confidence in conducting lab-based research, scientific and critical thinking, motivation in STEM education, and interest in pursuing graduate programs and the chances of being accepted (Russell et al., 2007; Adedokun et al., 2014; Brownell et al., 2015; Olimpo et al., 2016). Byars-Winston et al.'s (2016) report suggests that participation in undergraduate research is a central element in increasing the interest and persistence of college students in science careers. Hanauer et al. (2016) agreed, indicating that course-based research experience directly responds to the need to enhance retention and improve educational outcomes in STEM education. The lengthier the undergraduate research experience is and the higher level of consistency of holding a position in one research lab, the greater likelihood for STEM students to stay in their majors (Cooper et al., 2019). Students who have a positive lab environment and enjoy daily research tasks are less likely to have the intention of leaving (Cooper et al., 2019).

These findings apply to URM STEM students as well. Chang et al. (2014) conducted an analysis that involved survey records of 3,670 students at 217 institutions. The result suggests that participation in undergraduate research programs is the strongest predictor of the likelihood of URM STEM students' persistence—17.4 percentage points more likely to persist in STEM than those who had no undergraduate research experience. In terms of the technical aspects of undergraduate research, Ghee et al. (2016) mentioned that the use of bibliographic software, statistics software, and quantitative data analysis, are considered extremely beneficial. Below are some program examples.

An evaluation of the Futurebound Program, launched in Arizona, showed that students recognized undergraduate research opportunities, together with five other activities, as most helpful, and the Futurebound Program indeed raised retention rates (Reyes, 2011). By implementing a package of undergraduate research programs (i.e., research initiatives for science majors, research symposium, undergraduate research course credit) at a public Hispanic-Serving Institution (HSI), students' one-year retention rate increased from 69% to 77%, and STEM graduates' GPA (3.01 ± 0.05) remained significantly higher than that for all other majors at the institution (2.92 ± 0.03), during the program period (Carpi et al., 2013). The NSF-funded Research Experiences for Undergraduates (REU) program at New York City College of Technology has a focus on

Geoscience. It engages URM student participants in the year-round academic program (nine weeks of full-time internship in the summer, three weeks in the fall semester, and three weeks in the spring semester). Since 2008, REU has successfully supported 39 URM students to build up their careers in STEM fields, and the top-rated program features include structured preparation, student-centered mentorship, and diversity awareness across program networks (Blake et al., 2013).

The STEM-ENGINES program in the Chicago area embedded the research element into the transfer pathway. Students participated in part-time research at a community college during the academic year and in full-time research for eight to ten weeks at four-year institutions in summer sessions (Higgins et al., 2011). Students in the above-mentioned programs experienced gains in confidence during the transfer process. Nevertheless, it should be noted that participation in research opportunities does not guarantee successful transfer in all scenarios. For instance, the NanoSTEM Program at Binghamton University enrolled only high-quality STEM students who were already interested in pursuing a STEM career; however, two students of three ended up leaving the program (and the school) for medical reasons, and the third quit after eight weeks after not adjusting well to the university (Cott et al., 2016).

While undergraduate research is important in stimulating strong, positive personal and professional developments for URM STEM students, researchers found barriers regarding undergraduate research: Sens et al. (2017) noted that not enough research opportunities are available on campus; Pierszalowski et al. (2021) argued that it has not paid sufficient attention to diversity and inclusion; Mahatmya et al. (2017) surveyed students who were never engaged in research projects and found that lacking available mentors and the disproportionate assistantship employment was most frequently mentioned barriers.

Faculty and Peer Mentorships

Faculty Mentoring. Mentoring relationships are direct personal connections between STEM students and faculty members who are familiar with one or more aspects of the STEM department, its academics, and possible future careers (Ong et al., 2018; Xu, 2018). In acknowledgment of the many positive effects of faculty mentoring, many grants to colleges have as a condition that an undergraduate mentoring program is created to provide students with the support and guidance that they need to persist in STEM education (Olson & Labov, 2012).

Students in the Benjamin Banneker Scholars Program (BBSP) at a Midwestern historically Black college perceived that faculty mentoring was the biggest contributing factor to their persistence (Kendricks et al., 2013). Student surveys conducted by Kendrick et al. (2013) also indicate a strong correlation between students' academic success and their acceptance of mentoring as a positive experience in their learning. This relationship also enables underrepresented students to develop a sense of belonging and more positive scientific identities (Potvin & Hazari, 2013), which are key for retention and success. In a study conducted by Ghee et al. (2016), it was found that the specific quality of the mentor had the largest effect on the research skills index. Mandatory monthly advisory meetings, diverse instructions, and role models can all improve mentoring quality. On the other hand, low-quality (even harmful) mentoring is considered to be deceitful, sabotaging, exploitative, and/or harassing (Eby et al., 2000). For URM STEM students, quality mentorship, together with research experience, during junior and senior years were positively related to their scientific efficacy, identity, and values. Scientific identity and values even continued to be predictive of STEM career pathway persistence for up to four years after graduation (Estrada et al., 2018). Even informal supportive faculty members can be beneficial, including but not limited to joining faculty members for lunch, encountering them in the hallways, being known by name, and being taken care of in a close personal way (Nuñez & Yoshimi, 2017).

Some researchers observed that demographic similarities between faculty mentors and mentees (the so-called minority mentorship) increased the chance of positive mentoring experiences and academic success in STEM fields (Blake-Beard et al., 2011; Atkins et al., 2020). Developing mentorship with role models of similar cultural backgrounds would help provide culturally appropriate learning experiences for URM STEM students. It allows the construction of solid racial identity and ultimately improves the student’s learning experience. Adequately contributing to psychosocial development can enable a mentor to maintain a strong level of racial awareness and an appropriate approach to addressing complex racial situations as they may arise. The five progressive and complementary achievements through effective minority mentoring include gains in the mentee’s professional competence; increased mentee’s confidence and credibility; prevention of mentee’s derailment; powerful mentor sponsorship of the mentee; and mentor protection of the mentee in unfair or unjust situations, such as racial disparagement (Thomas, 2001). Moreover, minority mentoring is not a unidirectional transfer of information from the mentor to the mentee but tends to be reciprocal in nature (Higgins & Thomas, 2001; Rock & Garavan, 2006). However, Carroll and Barnes (2015) pointed out that often, there are not enough senior URM mentors in STEM disciplines to support the number of URM students needing mentorship. They proposed that cross-cultural mentoring would facilitate growth if the mentors adopted an infusing broaching style to address uncomfortable sensitive racial topics, instead of being colorblind or inconstant in their words and actions. This would certainly require a higher level of cultural competency and communication skills. However, cultural competency improves by receiving feedback from students on their mentoring styles and strategies. In addition, some researchers found weak or negligible support for a link between demographic similarity and the quality of mentorship (Eby, et al., 2007; Eby et al., 2013). In particular, when mentors and mentees have little contact, they tend not to know much of the degree to which they shared similar demographic characteristics. The observed and perceived similarity might, therefore, have only limited influence on the quality of mentorship early in the faculty mentoring relationship (Harrison et al., 2006).

Hernandez et al. (2017) suggested that shared values are more responsible for mentees perceiving higher mentorship quality than demographic match. Robnett et al. (2019) confirmed through an intersectional analysis showing that higher levels of instrumental mentoring's predictions for higher STEM self-efficacy are likely not moderated by ethnicity or gender. The amount of mentor-mentee contact is an important moderator of the relationship between demographic similarity and the quality of mentorship, and relationship satisfaction was the only moderator of the effect of perceived similarity. Other reports (Higgins & Kram, 2001; Hernandez et al., 2017; Robnett, et al., 2019; Atkins et al., 2020) have indicated that compared with weak mentor-mentee ties in informal faculty mentoring, mentoring with deeper contact (i.e., with higher interaction frequency, longer relationship duration, instrumental plus socioemotional mentoring) will bring greater positive effects, particularly in the context of socioemotional mentoring. In an immersive faculty mentoring program, where participants were required to live together in an Honors Dormitory and take part in mandatory meetings and research, minority student participants perceived that mentoring was the biggest contributing factor to their academic success. Mentors in physics and chemistry, mentors observed that corrective actions took additional time, and students showed the greatest improvement during the final five-week period of the semester (Kendricks et al., 2013). In another example described by Zaniewski and Reinholz (2016), faculty mentors provided holistically designed near-peer mentoring, which combined psychosocial support (most closely linked to identity formation and belonging) and academic support (to promote self-efficacy and thus belonging). Students received credits from the mentoring program and caring friendships with mentors. As a result, in the population served by the mentoring program, the first-to-second-year major persistence rate increased from 59% to 93%.

Peer Mentoring. Peer mentoring incorporates peer instruction, small-group activities, extra worksheets, practice tests, and other means of interaction between peers. Damkaci et al. (2017) found a 48% difference in persistence rates in STEM majors between students who participated in a peer-mentored lab course and those who did not. Peer-to-peer relationships function because they provide a combination of social, academic, and emotional support through shared experiences, intersectional identities, and working together toward mutual success (Ong et al., 2018). Musah and Ford (2017) argued that peer mentoring helps STEM students improve their grades in STEM courses and reach higher GPAs in subsequent semesters, although neither of these improvements appeared among transfer students who participated in the same program. Morganson et al. (2015) used the embeddedness theory (Mitchell et al., 2001) to incorporate the broader community context into their inquiry. They found that solidarity and friendship with peers studying in STEM majors, role models, and cultural icons were critical positive factors for building a sense of community embeddedness, while passion, challenge, skills, aptitude, concrete tasks, real-world applications, and intrinsic outcome expectations were positive aspects of major embeddedness.

The duration of mentoring matters. Anderson and Kim (2006) found that at the three-year point, all ethnic groups were almost equally likely to continue STEM enrollment. But following the third year, the progress of Black and Latinx students had slowed relative to that of their counterparts. Therefore, it is important to continue or develop peer mentorships in the junior and senior years of URM STEM students' undergraduate study. Those years are also the right moment to introduce peer mentoring programs for transfer students. Brown et al. (2016) focused on matching peer mentors and transfer students. All of the participants in their study successfully persisted to the following semester. The advantage of the studied program lay in its opt-out rather than opt-in policy: all new students were assigned a mentor and had to apply to withdraw from mentoring if they wished. In that way, students who did not feel comfortable asking for help or who did not initially perceive a need for mentoring could have access to a mentor when the need arose (Brown et al., 2016).

To improve outcomes, Wilson, Holmes, et al. (2012) introduced a hierarchical mentoring strategy requiring that mentees must mentor other peers themselves to reinforce the consistency of transformation of survival skills among minority students. They also indicated that mentees who entered the mentoring program at the sophomore level had a better chance of completing a STEM degree than those who entered it at the freshman level. In addition to that, Packard (2016) proposed that mentoring program should be developed based on outcome planning. This intentional mentoring links mentors' actions, approaches, goals, and indicators to improve the effectiveness of peer mentoring.

A combination of peer mentoring and faculty mentoring has been found to effectively support URM STEM students' persistence. At the University of Wisconsin-Whitewater, the enrollment persistence to year two for student participants (96%) was significantly higher than the expected retention rate (72%) for URM students in the university and even exceeded the university's non-URM rate of persistence to year two (80%; Lisberg & Woods, 2018). Some programs even provided multiple peer and faculty mentors for each student participant over multiple years. In the Significant Opportunities in Atmospheric Research and Science (SOARS) program, managed by the University Corporation for Atmospheric Research (UCAR), mentors had functions in different aspects of student life, including research, communication, writing, computation, and community development. Students joined their mentors over the summer vacations throughout their undergraduate years. This repeat mentoring experience was whole-person-oriented and resulted in sending over 90% of mentees to graduate schools (Feder, 2019). However, it was uncommon to have continuity in this type of mentoring relationship (Robnett et al., 2019).

Beyond the interventions mentioned above (i.e., financial support, course-based undergraduate research, and faculty and peer mentorships), institutional arrangements, including institutional control over key first-year courses' teaching quality (Xu, 2018), hands-on workshops for professional development (Reyes, 2011; Tupper et al., 2010), and networking events (e.g., onsite company visits, job interviews, and career fairs; Tupper et al., 2010) have also been discussed in previous studies. However, even after being equipped with these activities and resources, institutional barriers and resource disparities still exist among URM STEM students (Estrada et al., 2016).

The Theory-Based Evaluation Approach

Over the decades, scholars have infused insights into program evaluation. These insights help evaluation evolve from an applied social research method to a way of thinking, introduce it into the wider social life, and enrich it with alternative models and approaches. Evaluation theories have developed from positivist randomized control experiments (Campbell & Stanley, 1966) to much more plural perspectives (Guba & Lincoln, 1981; Dahler-Larsen, 2011; Schwandt, 2015). The concern shifts from the validity of methodologies to utilizable implications and social value. Theory-based evaluation falls more in the middle, providing a valuable theoretical framework for evaluators to assess both organizational actions and their theory of change.

Theory-Based Evaluation

Theory-based Evaluation (TBE) is an evaluation approach that emerged in late 1970s, responding to the shortage of randomized control trials (RCT) in evaluation practices (Weiss, 1997). TBE argues that the traditional experimental methods lack focus on the entangled relationships between the intervention and related variables. The outcomes from such research-type evaluations often generate oversimplified or even distorted understandings of interventions and their impact (Chen & Rossi, 1983, 1987). This default hypothesis that the intervention leads to linear and short-term gains indicates a lack of understanding of the real trajectory of solving social problems (Woolcock, 2009). The overemphasis on the internal validity of the methodologies in evaluation impairs the ability to apply the result to real-world practice, let alone informing program improvement and policy development (Weiss, 1997).

Although TBE has no obvious philosophy statement, research shows that the most frequently occurring motive for selecting a TBE approach is ideological (Coryn et al., 2011). It is usually considered to be a critical realism approach that addresses ontological realism, epistemological relativism, and judgmental rationality in practice (Bhaskar, 2008; Brousselle & Buregeya, 2018). Critical realism is one of the most common forms of post-positivism, which is situated between positivism and constructivism or relativism (Bhaskar et al., 1998). TBE scholars believe that truth exists independent of individual subjectivity and experience. The real structures and the actual causal pathways of object knowledge cannot always be directly observed but they can be inferred through observable events and results. Only when we understand the structure of logical relationships and causal chains, which is a necessary precondition for what can be observed, can we understand how and why outcomes occur. Therefore, it is important for evaluators to intentionally and rationally explore, construct, and examine the program theory based on existing knowledge in order to determine the worth and merit of the evaluand (Brousselle & Buregeya, 2018). In terms of epistemology, the TBE approach recognizes that reality can only be understood fractionally within a particular context, so it respects the complexity in the system that nests the intervention. The evaluation process, therefore, incorporates subjective values, perceptions, judgments of reality, and contextual influences into account to some extent. Meanwhile, one should be aware of cognitive limitations and be open to other explanations and unexpected evaluation results that may exist (Pawson & Tilley, 1997; Brousselle & Buregeya, 2018). Regarding methodology, TBE's stance is relatively neutral. Any research method that contributes to the understanding of the program theory can be used in the evaluation process—provided that the method's reliability and validity requirements are met. Therefore, TBE can accommodate a variety of social science research methods, data collection methods, and data analysis methods. For TBE, quantitative, qualitative, and mixed methods have their own merits and are suitable for different evaluation needs and purposes. In TBE practices, the use of mixed methods is justified and applied under a program theory framework, which is defined as a set of explicit and/or implicit, prescriptive, and descriptive assumptions held by stakeholders about what actions are required to solve a social problem and why the problem will respond to these actions (Chen, 2006; Coryn et al., 2011; Brousselle & Buregeya, 2018).

In evaluation practice, a logic model is commonly used. TBE can also utilize logic models to specify the program theory, especially when program staff members are also engaged in the construction of the program theory (Kaplan & Garrett, 2005). However, the process recorded in the logic model is simplified. The "arrows" between every two parts do not guarantee actual causal connections, and the real mechanism is still hidden behind the model. TBE does not aim to report expected linkages between resources, courses of action, and outcomes, but to provide a validated model for explanation and prediction.

TBE assists evaluators to investigate the impact of the evaluand (i.e., program, project, policy, etc.) and the process by which impacts occur. TBE is sometimes referred to as program-theory evaluation or theory-guided evaluation (Coryn, et al., 2011) and it has become an umbrella approach that contains all evaluation strategies that explicitly integrate and use program theories generated from stakeholders, social sciences, or any other sources in conceptualizing, designing, and conducting an evaluation (Coryn et al., 2011). It emphasizes the importance of unpacking the “black-box evaluation” (Astbury et al., 2010; Chen & Rossi, 1987). The metaphor “black box” has been used to refer to the invisible inner logic and program theories of an evaluand, indicating that the process between inputs (e.g., resources, activities, knowledge) and outputs/outcomes is usually unclear (Astbury et al., 2010). Instead of merely asking whether and to what extent the evaluand achieves its goals, providing information about the performance and outcomes, or making judgments on the process and results of a program, TBE tries to unpack the “black box” by explaining how and why the program achieves the results in order to address both scientific credibility and practical worth when serving stakeholders’ evaluation needs (Weiss, 1997; Birckmayer & Weiss, 2000; Rogers & Weiss, 2007). Through this approach, evaluators can help reason the success or the failure. Sometimes the problem may come from program implementation (Chen, 1990), the changing context (Pawson & Tilley, 2004), and the complexity in the system (Westhorp, 2012, 2013, 2014). By understanding the process and factors, the evaluation leads to more targeted recommendations to the evaluand.

Theory-driven evaluation, realist evaluation, contribution analysis, and logic analysis are TBE frameworks that are mature in their design and application (Coryn et al., 2011; Brousselle & Buregeya, 2018). Since the latter two are relatively new and have a lot in common with the former two regarding concepts and procedures, this review will focus on the former two. Theory-driven was introduced in a series of manuscripts jointly published by Chen and Rossi in the 1980s (1980, 1983, 1987, 1989), in which they discussed the concept, method, and limitations of theory-driven evaluation. These discussions were eventually integrated into a theoretical framework by Chen in the book published in 1990. In the following years, Chen continued to horn his opinions on the methodology and its application in different evaluation stages (Chen, 1989, 1997, 2006, 2014). Realist evaluation, proposed by Pawson and Tilley (1997), has a strong focus on the underlying mechanism (i.e., the causal theory) of the evaluand by engaging specific contextual factors required for the mechanism to function. The key components in realist evaluation are the context within which activities occur (C); the mechanism, which is how human actors respond to interventions (M); the outcome created by the mechanism (O). The equation C + M = O (CMO) stands for a hypothesis of the interactions among these components and tells what works for whom in what circumstances (Pawson & Tilley, 1997, p.58). There are usually multiple CMOs contained in one intervention. Realist evaluation does not avoid the issue of complexity but decomposes the elements of complexity and understands it through hypothesis testing (Westhorp, 2012).

The Theory-Driven Evaluation Framework

Chen (1990) defined program theory as "a specification of what must be done to achieve the desired goals, what other important impacts may also be anticipated, and how these goals and impacts would be generated" (p.43). This definition reflects the TBE scholars' critical realism positioning. In theory-driven evaluation framework, program theory consists of two models: an action model and a change model. The action model describes a systematic plan of the organizational structure, partnerships, and organizational activities that deliver the intervention, plus the influence of internal ecological contexts and external resources. The change model is the logic that describes how interventions will finally achieve the goals and outcomes, either wanted or unwanted. These two models are connected by the program implementation. Thus, this program theory framework demonstrates that the program runs from resources coming into action model built up, intervention implemented by the organization, and the appearance of the theory of change. The consideration of the theory of change should influence the previous steps but it will not be completely formed until the end of the program because the explanation of unwanted outcomes is also part of the theory of change (Chen & Rossi, 1980; Chen, 1990). A program theory is generally the stakeholders' theory. However, if stakeholders do not systematically clean up and map out the program theories, which is usually the case, the evaluators will need to review literature, analyze program documents and records, and facilitate interviews with stakeholders to clarify or even develop their theories of change before they can empirically assess the effectiveness of the program theory.

The theory-driven evaluation framework claims that the program system synthesizes reductionism and fluid complexity, which is assumed to be stable in the evaluation period and reduce the change caused by time and environment (Chen, 2012). It argues that the environment surely creates uncertainties that push a program to make changes, but a program can and should have proactive measures to reduce uncertainties and maintain some level of stability for performing its functions. However, this assumption is usually violated in a dynamic open system, within which uncertainties are constantly generated by interactions among internal and external program elements. Furthermore, the theory-driven evaluation framework explains how it works when evaluators analyze the theory of change, but it becomes vague regarding the knowledge transfer from the analysis to the real-world practice, especially when the environment changes. As Chen (2012) points out, an effective and efficient program theory does not guarantee a good intervention if it is unsuitable for an organization to implement it. In addition, theory-driven evaluators deem that the evaluation should have a focus on learning from the evaluation that can be generalized to other programs (Sridharan & Nakaima, 2012). Therefore, this approach does not aim only to achieve the internal validity of research and evaluation designs, but also provide a general framework to deal with internal, external, construct, and statistical conclusion validity (Chen & Rossi, 1987). However, it turns back to relying on the sampling strategies and the statistical modeling that limit its strengths in generalization.

The Realist Evaluation Framework

In the decades following the presentation of the theory-driven evaluation framework, investigators sought to explore various approaches under the TBE umbrella, including theory of action, theory of change, program logic, logical frameworks, outcome hierarchies, and realist evaluation (Coryn et al., 2011). As a branch of TBE, the realist evaluation framework became better known in the late 1990s and early in the 21st century. Although new generations of evaluation approaches have taken some of the spotlight in the recent decade, continuous practices and discussions around this framework are pushing it toward another surge, especially in Australia and European countries (Betts, 2013; Holma & Kontinen, 2011; Jagosh et al., 2016; Magnin et al., 2018; Mansoor, 2003; Manzano, 2016; Pommier et al., 2010; Westhorp, 2012, 2013, 2014). Realist evaluation moves forward based on the theory-driven evaluation framework in regard to its response to the three advanced, context-based inquiries in evaluation: deal with complexity, translate the evaluation results into practices, and apply an intervention to other contexts.

Pawson and Tilley (1997) saw the purpose of evaluation as informing the development of policy and practice. Instead of providing a binary evaluation judgement as either a “success” or a “failure”, they developed the realist evaluation framework to determine what works for whom in what circumstances, which they extended to what works for whom in what circumstances, in what respects, and how in their later work (Pawson & Tilley, 2004). Pawson (2013) argued that “[a] program works because of the action of some underlying mechanisms, which only comes into operation in [a] particular context” (p. 22). Thus, realist evaluation is an applied method that explains the process of how interventions or program activities lead to the outcomes through the interactions with complex components in the context. This process is called a mechanism. Realist evaluators usually consider an evaluation to be a form of research that has a specific focus on a complex intervention because it encourages evaluators to dissect the complex intervention and find pathways toward both expected and unexpected outcomes (Astbury et al., 2010; Pawson & Tilley, 1997). Some scholars have criticized this, saying that it goes beyond evaluators’ capability as the role of evaluators is to determine whether programs work, not to explain how they work (Scriven, 1998). Stufflebeam and Shinkfield (2007) added that finding possible paths is usually not feasible, and failed or misrepresented attempts can be highly counterproductive.

The realist evaluation’s critical realism ontological foundation responds to the latter comments. Critical realism addresses both observable and unobservable contextual situations and tries to offer solutions of identified problems (Bhaskar et al., 1998; Porter et al., 2017). Such positioning asserts that the adequate explanation of social phenomena should consider the involvement of human factors. It is necessary to identify how individuals interacted through the mechanisms contained in the intervention and embedded in inter-connected and multi-layered context (Porter et al., 2017; Jamil, 2018). Although it recognizes the existence of a mind-independent external reality that science can study, it asserts that human beings will never measure and document true reality with certainty because we experience and understand it via our senses and brains (Archer et al., 2013). Therefore, critical realism holds that all observation is fallible, and that all theory is revisable. Human beings' understanding will always be partial and provisional, and evaluators should not be over-ambitious to map out all mechanisms in a program at one point in time. However, the limitation in the capability of knowing does not suggest that we are forever blind to the mechanisms and consequences of interventions. All the trials and errors are valuable in terms of theory refinement and modification, including those failed or misrepresented attempts. It is important to keep in mind that a program has different levels of success with different participants in different contexts—and even in the same context at different times (Greenhalgh et al., 2015).

With an intention to inform policy and practice, a realist evaluation differs from other types of TBE (e.g., theory of change and contribution analysis) that pay attention to the overall program theories. It focuses on the particulars of specific measures in specific places relating to specific stakeholders (Pawson & Tilley, 1997), expecting measures to be varying in their impact depending on the conditions in which they are introduced. Accordingly, this increases its ability to transfer the knowledge pieces into particular policy and practices. Whereas science is concerned with understanding regularities (context—mechanism—regularity), program evaluation involves understanding how regularities change (context—mechanism—outcome configurations). In this sense, realist evaluators are pragmatists as well. As Tilley (2000) mentioned, realists see benefits in limited applications of interventions in order to understand their effects before making decisions about their extension.

Realist evaluation is the process of proposing the hypotheses, analyzing the interaction, modifying the equation, and generating a new hypothesis accordingly. This conceptual framework connects theories and practices. There is no specific ending point as there is no absolute moment at which we reach the truth. In particular, the realist evaluation approach advocates that one evaluation should learn from another, supposing that evaluation can learn lessons from diverse programs by operating at the middle range. A sequence of realist evaluations will lead to more powerful context—mechanism—outcome configurations (CMOCs), although this has not yet been proven (Mansoor, 2003; Pawson & Tilley, 1997).

As for the broader application of the intervention, the realist evaluation tends to maximize external validity by learning across policy, practice, and organizational boundaries (Pawson & Tilley, 2004; Tilley, 2000). The conceptual abstraction allows the creation of a common language that draws out similarities across different interventions (Pawson, 2013). One metaphor is that "programs are theories incarnate" (Pawson & Tilley, 2004, p. 256). There are multiple sources of theories, including stakeholders, policy makers, social sciences, evaluators' observations, and even common sense (Mansoor, 2003; Westhorp, 2014). Some scholars have suggested that the realist approach should be applied after finishing a Campbell-style systematic review of social science theories and topic-related practices (Van der Knaap et al., 2008). In that, the more thorough the background knowledge is, the more solid the proposed hypothesis could be. A realist evaluation is sensitive to diversity and change in program delivery and development, but this does not mean that the intervention could not be implemented to other circumstances at all. The CMOCs help understand how an intervention works with the mixed components. Thus, the comparisons among CMOCs clarify which ingredients change and how this will influence the outcomes.

Four-Step Process for Undertaking TBE

Although there are slight differences in the operation process across TBE frameworks (White, 2009; Coryn et al., 2011; Brousselle & Buregeya, 2018), in general, evaluation applying TBE usually goes through the following four steps:

Establish a Preliminary Program Theory. The program logic and the expected causal chain should be identified to guide evaluation. This can be realized through communication with program leadership, administrators, implementers, and experts on related issues to understand the resources invested in the program, the action theory, and the implementation process; collecting and reviewing program documents (e.g., memoranda, grant proposal, meeting minutes, logic model) to understand the nature and the context of the evaluand.

Take Contextual Factors into Account to Hone the Program Theory. In this step, evaluators may collect and review academic and non-academic literature related to the program-related issue and understand the key factors that have a potential impact on the process. Then, evaluators will identify contextual factors (e.g., the socio-economic and political setting, local demographic characteristics, human factors in program implementation, program management issues, etc.) and place the causal chain in the specific context. If there are multiple sets of contextual factors that influence the mechanism, consider establishing separate causal pathways thereby linking the different outcomes that the program may end up with where the mechanism works.

Design An Evaluation Plan and Conduct Theory-Based Data Collection. The evaluation plan integrates the causal pathways, their corresponding theoretical assumptions, and data collection methods. The data types and the data collection process are determined based on the needs. Quantitative and/or qualitative methods are applied accordingly.

Examine the Program Theory and Refine It According to the Results. Evaluators can bring the results to stakeholders and discuss the results with the program's leadership, administrators, implementers, and experts on related issues, listening to their opinions and questions. Evaluators also pay attention to unwanted or unexpected outcomes and adjust the causal chains based on the results and the context. The last two steps can be repeated to improve the explanatory power and accuracy of the program theory.

In all, TBE requires the evaluand (1) to have a guiding theory and an action logic; (2) to have entered a relatively stable implementation stage and be able to provide a certain amount of data supporting for theory examination; (3) to request an evaluation that supports the theory-informed decision-making. TBE has a clear focus, and the evaluation results provide information that helps understand the theory of change in addition to the value judgment. TBE helps stakeholders clarify how and under what conditions program resources are transformed into outcomes and provides evidence to inform the generalization and application of intervention in a broader scope of contexts.

Critical Race Theory Lens

The quantitative analysis of factors affecting the retention, persistence, and graduation of URM STEM students tells only part of the story. It does not tell how power dynamics and individual positionality influence students' experience. These aspects may, instead, be investigated through studies employing a Critical Race Theory (CRT) lens. Studies that employ CRT offer a perspective that explores the voices of historically marginalized individuals. CRT promotes insight into race- and ethnicity-related social issues that challenge the ways that racial power is constructed. This is one of the core concerns of the current study, focusing on URM students in STEM majors.

CRT is a powerful explanatory tool for the sustained inequity that people of color experience in education (Ladson-Billings, 1998). Bell (2003) asserted that racism is hegemonic and exists permanently in educational institutions, overtly or covertly. Racial microaggressions, including interpersonal microaggressions (verbal and nonverbal racial affronts), racial jokes, and institutional microaggressions (through university and local culture, organizations, informal rules, implicit protocols, and institutional memories) create a negative campus racial climate that exacerbates student psychological distress (Sue et al., 2008; Yosso et al., 2009). Members of different racial groups have differential access to quality education. Hence, an important objective of CRT is to examine social and cultural phenomena in diverse settings through critical analyses of race, law, and power to show the existing and historical nature of racism at the institutional level (Bell, 2009). By analyzing the continuance of racism and white privilege in education, CRT can address systematic ways in which students of color are disadvantaged. It highlights the profound patterns of exclusion that exist in the existing educational system. CRT is thus a vital approach to the study of higher education for resolving racial inequality (Hiraldo, 2010).

CRT scholarship emphasizes personal narratives and advocates the empowerment of marginalized individuals to share their experiences themselves: to perform counter-storytelling (Delgado & Stefancic, 2017; Solórzano et al., 1998; Yosso, 2013). It indicates that persons of color do not share a single voice in common. Narrative analysis serves to illuminate their individual experiences of racial oppression (Solórzano & Yosso, 2001). Scholars take personal stories and experienced knowledge as evidence, thereby challenging purely quantitative approaches to the documentation of inequity or discrimination (Dixson & Rousseau, 2014). CRT insists on a critique of liberalism—racism requires sweeping changes, but liberalism has no mechanism for such change. CRT recognized an aggressive and race-conscious approach to social transformation (Delgado & Stefancic, 1993). CRT employs radical activism to explore and challenge the prevalence of racial inequality in society, because institutionalized racism is steeped in existing social structures and norms, resulting in a range of levels of access to services, opportunities, and power for those of different races (Jones, 2002).

Harper (2012) indicated that “studying race without racism is unlikely to lead to racial equity and more complete understandings of underrepresented populations in postsecondary contexts” (p. 15). Scholars have proposed that in examinations of student development, it is insufficient to include students of color in research samples. CRT theorists call for a paradigm shift, involving efforts to seek out and to discover how social forces and inequities influence self-authorship, including how students view themselves, what they believe, and how they construct relationships with others (Hernández, 2016). As noted, many programs provide support in a general sense while often ignoring individual needs and perspectives relating to cultural characteristics (Byars-Winston et al., 2016). Transfer students face challenges related to gender, race, and socio-economic characteristics. They may encounter barriers in the transfer pathway that their counterparts do not. CRT should be utilized to probe these aspects and help modify the design of interventions by taking transfer students' needs and perceptions into account.

In all, CRT offers a framework that engages the voices of historically marginalized individuals to investigate inequity, reveal the obvious and hidden oppressions and subordination in educational practices, and promote insight into race- and ethnicity-related barriers and challenges for URM STEM students (Ladson-Billings, 1998).

Chapter Summary

Previous studies have suggested that URM STEM students face both general and unique challenges and barriers on account of the intertwined effects of demographic, academic, and social-cognitive factors. Four-year and two-year institutions have endeavored to narrow the gap by supporting URM STEM students through financial support, course-based undergraduate research connections, faculty and peer mentorships, and many other interventions. Research and evaluations have shown the effectiveness of these interventions in some contexts. Operating for 15 consecutive years, these interventions have been implemented at LS-NSSA as well. There is a lack of understanding of the metrics and the effectiveness regarding key factors. It is worth examining how and why this program works for the students it serves.

The theory-based evaluation approach is proposed to be a productive way to approach the evaluation. This approach facilitates examination of the program's theory of change. The mechanisms that drive URM STEM students' retention and graduation are explored via social science theories as well as examination of the target program's theory of change. To address systemic inequity in the educational experiences of URM STEM students, it is necessary to identify and evaluate culturally appropriate, evidence-based interventions. Theory-based evaluation, in combination with critical race theory, can serve as a framework to investigate the research questions and ground the methods used in the current study.

CHAPTER III. METHODS AND PROCEDURES

An explanatory sequential mixed methods design was applied to answer the research questions. Quantitative methods were utilized in two sub-studies. Sub-study 1 investigated research question 1: What are the factors that influence the retention and graduation of LS-NSSA's target population—URM STEM undergraduate students? Sub-study 2 investigated research question 2: How do LS-NSSA participants perform on key metrics compared with their non-participant counterparts? Sub-study 3 utilized qualitative methods to examine research question 3: How do participants perceive the importance of LS-NSSA's research mentorship experience? This chapter describes the methods and procedures for three sub-studies.

The Mixed Methods Design

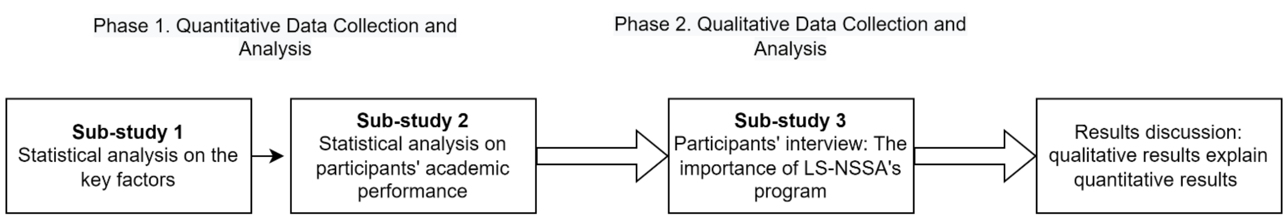
Tashakkori and Teddlie (1998, 2021) provided a comprehensive overview of the mixed methods approach. They indicated that mixed methods research "combines the qualitative and quantitative approaches into the methodology of a single or multiphase study (Tashakkori & Teddlie, 1998, p.18)". The application of different methods realizes the methodological triangulation that allows a multi-dimensional understanding of a phenomenon (Denzin, 2011). Based on the definition and the discussion of triangulation techniques, scholars classified mixed methods designs into several types (Creswell, 1999, 2010; Plano Clark & Creswell, 2014; Creswell & Plano Clark, 2018) and identified three core designs: the convergent design, the explanatory sequential design, and the exploratory sequential design (J. W. Creswell & J. D. Creswell, 2017). The convergent mixed methods design is a single-phase design in which the same or parallel variables are used to collect both qualitative responses and quantitative data. Different data types are integrated into the data analysis or the interpretation process to yield jointly displayable results (J. W. Creswell & J. D. Creswell, 2017). The explanatory sequential design contains a two-phase data collection procedure in which quantitative and qualitative data are separately collected and analyzed in two distinct phases—the quantitative procedure is conducted first, followed by the qualitative procedure. Quantitative results inform the plan of the qualitative phase, including research question development, sampling arrangements, and data interpretation (J. W. Creswell & J. D. Creswell, 2017). The exploratory sequential design consists of three phases. It starts from a qualitative study, whose result is used to guide the feature identification, variable selection, measurement instrument development, and experiment design (the second phase). Then, a quantitative study is conducted to test the model developed in the second phase (J. W. Creswell & J. D. Creswell, 2017). In a mixed methods study, qualitative and quantitative data are integrated throughout the data collection, analysis, and interpretation procedures to draw contextualized perspectives (Johnson et al., 2007).

The current study applied an explanatory sequential design to investigate the target program's theory of change. Quantitative data was first collected and analyzed, identifying the factors influencing URM STEM students' retention and degree attainment (sub-study 1) and testing LS-NSSA's participants' performance regarding these factors (sub-study 2). Sub-study 2 was informed by sub-study 1 in that sub-study 1 identified the variables chosen to be tested in sub-study 2. These two sub-studies led to the need for a further explanation of what students experienced at LS-NSSA that resulted in the observed performance. Therefore, the follow-up qualitative data collection and analysis (sub-study 3) complemented the quantitative inquiry. In this two-phase process (see

Figure 1), quantitative data provided a general picture of the outcomes and changes regarding URM STEM students' outcomes; qualitative findings helped explain how these changes were linked to each individual's program experience and why these changes occurred. Specifically, this design assessed quantitative (quantified) academic data collected from the Office of Undergraduate Education, University of Minnesota, and LS-NSSA's program documents; qualitative data were collected from in-depth exit interviews with LS-NSSA's research mentorship program participants. The analytical stage involved statistical and content analysis of the data.

Figure 1

Two-Phase Explanatory Mixed Methods Design of This Study

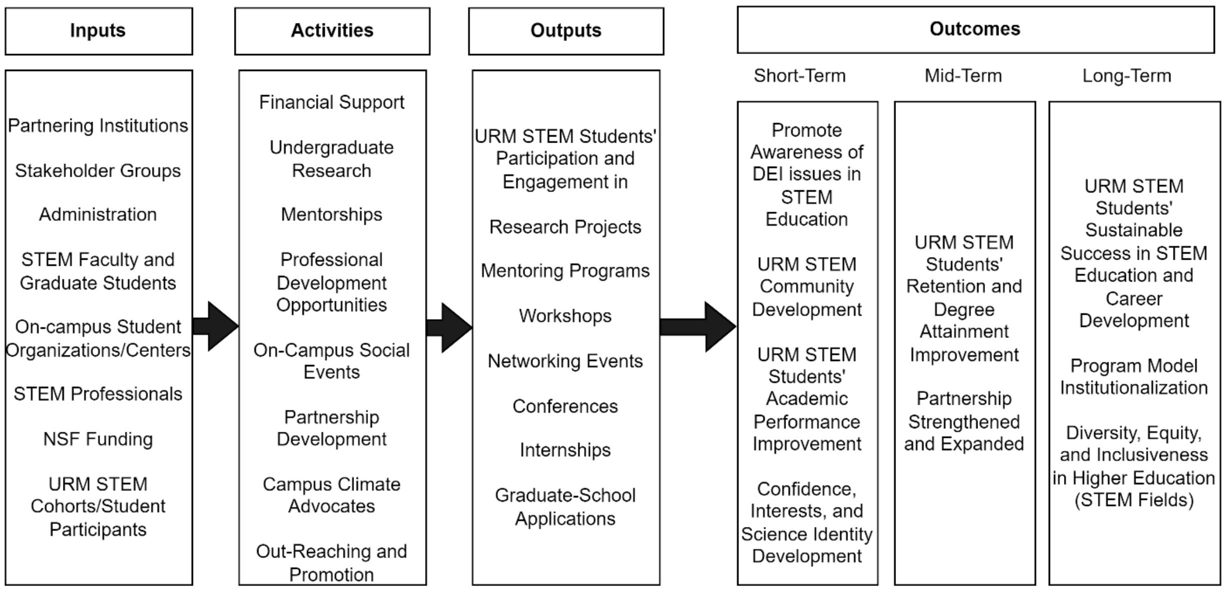


The Theoretical Framework: LS-NSSA's Theory of Change

In a theory-based evaluation, the target program's theory of change is layered on top of every evaluation step, which informs the conceptual framework of each sub-study. To understand LS-NSSA's theory of change, the author reviewed LS-NSSA's program documents (e.g., the grant proposal, the goal statements, the logic model, the external and internal evaluation plan, the annual reports, and the previous internal evaluation research summary). The logic model and the evaluation plan contain the latest information regarding the program design, developed by the program director with support from the external evaluation team. The logic model illustrates the implementation-level information—the process between inputs, outputs, and outcomes—that reflects the strategic-level theory of change (see Figure 2 for the simplified version).

Figure 2

LS-NSSA's Simplified Logic Model for Phase III (2017-2022)



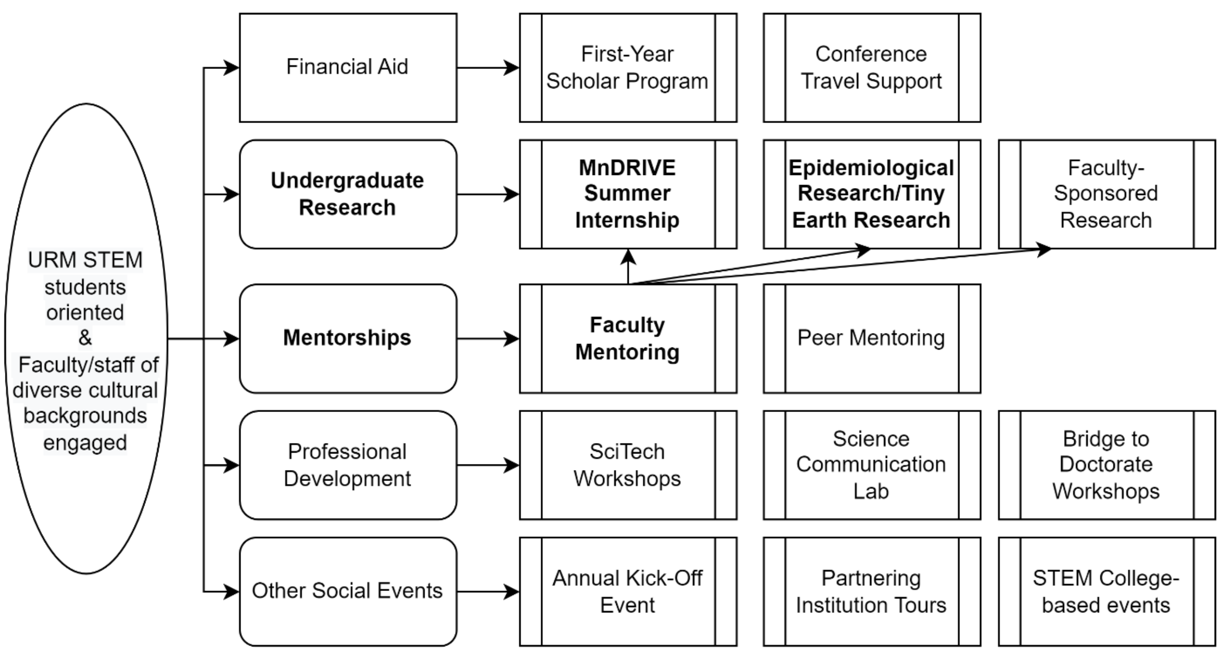
Note. This Logic Model is refined and simplified from the original logic model attached to LS-NSSA's 2017-2022 NSF grant proposal (LS-NSSA, 2016). This logic model integrated the designs from the program's Goal 1 (attain 756 underrepresented minority students earning STEM bachelor's degrees annually at Alliance institutions by 2022) and Goal 3 (foster institutional change toward greater diversity and inclusion on Alliance campuses), which are relevant to the internal evaluation's scope of work. The program's Goal 2 (deepen Alliance collaborations to improve students' pathways through all stages of transfer from community colleges to degrees at four-year institutions) has been primarily evaluated through external evaluation studies. See Appendix A for the entire logic model.

LS-NSSA, with NSF's funding, intends to provide different types of programs and services to support the target population's needs in various ways (see Figure 3). These programs and services connect program staff, stakeholder groups, administrative personnel, partnering institutions, STEM professionals, on-campus organizations, and

DEI (i.e., diversity, equity, and inclusiveness) initiatives with URM STEM students. They convert time, effort, and resources into students' participation and engagement. In the short term (i.e., in one to three years), the programs and services aim to advocate for a culturally diverse campus climate, increase URM STEM students' on-campus engagement, improve URM STEM students' academic performance, promote community development among URM STEM students and professionals, and strengthen URM STEM students' intention of retention. In the medium term (i.e., in three to five years), URM STEM students' retention rate and degree attainment are expected to increase because of increased confidence, identity, and commitment to the STEM field. In the long term (i.e., over five years), the program expects to witness sustaining success in URM STEM students' academic advancement and career development in STEM disciplines.

Figure 3

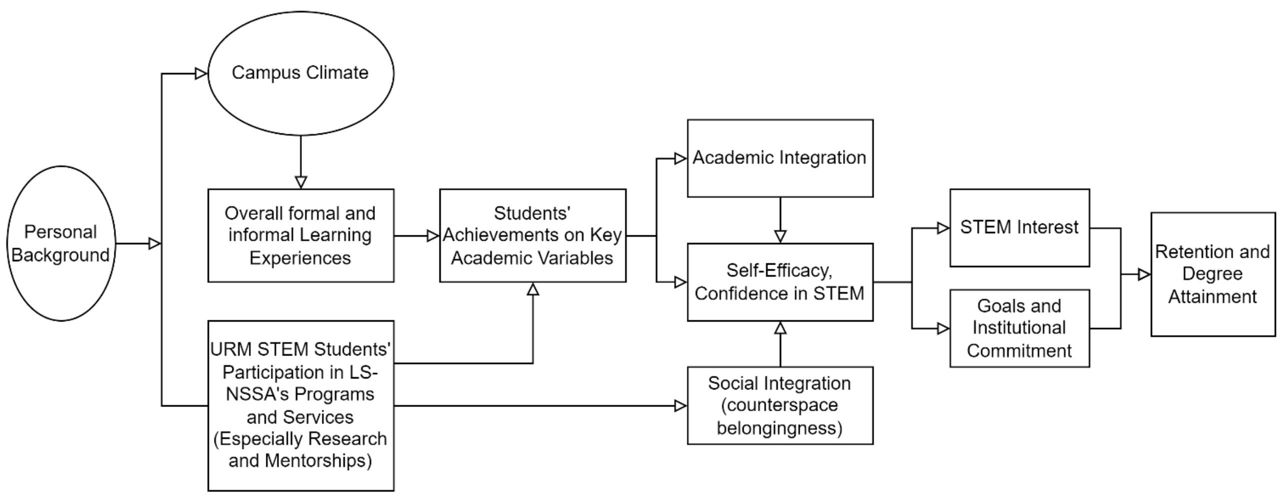
The Structure of LS-NSSA's Programs and Services



However, the logic model does not elucidate the key variables between URM STEM students' participation and their retention and degree attainment. There is a need to identify the elements that predict the outcomes of students' participation. Bandura's (1977) self-efficacy theory, Tinto's (1975, 1997) theory on STEM students' drop-off, and other studies summarized in Chapter II help shape the framework. The sub-studies aim to examine LS-NSSA's influence on students' academic achievement and social-cognitive development.

Figure 4

LS-NSSA's Theory of Change



Note. The theory of change is adapted from Tinto's (1997) model by integrating LS-NSSA's implementation strategies indicated in the logic model and other program documents.

Sub-Study 1 Methods

Sub-study 1 focused on academic factors that affect URM STEM students' retention and degree attainment by controlling demographic characteristics.

Definitions

STEM students were defined as students who declared a STEM major at the university using NSF's (2015) STEM Classification of Instructional Programs (also known as CIP Codes). The URM population was characterized as Black/African American, Hispanic/Latino, American Indian and Alaska native, native Hawaiian and other Pacific islander, and those of two or more races with at least one being underrepresented. Traditional students, also called new-high-school students, were the ones enrolled in the university right after graduating from high school or had no experience in secondary education. Transfer students were the ones who transferred from other higher educational institutions, two-year or four-year, to the university. The intra-university transfer students transferring from one campus to another were counted as traditional students because their experiences were like other new-high-school students compared with those transferred from other institutions. The three-year window was framed for transfer students because, on average, transfer students with an associate degree took 0.7 years longer to complete a bachelor's degree than those who entered the four-year institutions directly from high school. For those who entered college at an older age, the gap increased to 0.9 years (Shapiro et al., 2016). The data dictionary (see Table 1) indicates the variables contained in the dataset.

Table 1

Data Dictionary: Explanation of Variables

Variable Name	Description
Degree attainment status	Whether or not students graduated within the three-year window or four-year window*.
Likelihood of retention	The likelihood of students registering in the following academic year within the three-year window or four-year window.
Total transfer credits**	The total number of credits transferred from previous institution to the receiving institution.
Full-time status	Whether or not students register to be full-time students at entry.
First-semester GPA	Students' GPA at the end of the first semester.
First-semester credit-unit pass rate	Calculated by dividing the number of passed credits by the total number of credits registered in the first semester.
Stop-out records	The number of academic years that students have records of stop-out for one or more semesters within the year.
Gender	Male/female

Ethnicity	Students' recorded ethnicity.
Pell grant eligibility	Whether or not students are eligible for Pell Grant at entry.
Age	Students' recorded age at entry.
Housing status	Whether or not students live on-campus at entry.
Home location	Students' permeant residency status.
First-generation status	Whether or not students are the first person in their immediate family to attend college.

Note. * The three-year window refers to the common time frame for transfer students; the four-year window refers to the common time frame for students who arrived directly after high school to complete a bachelor's degree, described here as traditional students.

** Traditional students' transfer credits are college credits earned while enrolled in high school.

Samples

The dataset came from the Office of Undergraduate Education, the University of Minnesota (with IRB approval), the headquarter of LS-NSSA, which owns the majority of URM STEM students in the Alliance. The dataset contains seven cohorts enrolled between the academic year 2009-2010 to the academic year 2015-2016.

Across these cohorts, 22,955 STEM students were enrolled in three campuses that were LS-NSSA members during these years. Among these students, 1,720 (7%, 1,187 new-high-school students, and 533 transfer students) belonged to the URM population. In general, the URM population had a higher level of Pell grant eligibility and being first-generation college students in their households, especially for transfer students. URM students who were transferring from two-year institutions had a higher average age at entry (i.e., 25 years old) and mostly lived off-campus. Transfer students tended to have fewer passed credits in the first semester (11 units vs. 14 units) and a lower average first-semester GPA (2.7 vs. 2.9). Within the six-year window, transfer students had an average of almost 1.6 years with stop-out records. It was 0.1 shorter than traditional URM STEM students (see Tables 2 and 3).

Table 2

Descriptive Statistics for URM STEM Traditional and Transfer Students

Variable	Traditional URM STEM		Transfer URM STEM	
	n	%	n	%
Gender				
Male	598	50%	339	64%
Female	589	50%	194	36%
Race/Ethnicity				
Black/African American	420	35%	302	57%
Hispanic/Latino	424	36%	110	21%
American Indian/Alaska Native	343	29%	121	23%
Native Hawaiian/Other Pacific Islander	0	0%	0	0%
Pell Eligibility				
Pell	489	41%	353	66%
Non-Pell	698	59%	180	34%
First-Generation Status				
First-Generation	716	60%	352	66%
Non-First Generation	471	40%	181	34%
Graduation Status				
Graduated within the 4- or 3-Year Window	676	57%	351	66%
Not Graduated within the 4- or 3-Year Window	511	43%	182	34%
Housing Status				
Live On-Campus in the First Semester	531	45%	14	3%
Live Off-Campus in the First Semester	656	55%	519	97%
Registration Status				
Full-Time	1,186	100%	494	93%
Part-Time	1	0%	39	7%
Home Location				
Resident of MN Metro Area	638	54%	382	72%
Resident of Greater State Area	261	22%	75	14%
Resident of Reciprocal States	116	10%	24	5%
Resident of Other US States	169	14%	46	9%
Total	**1,187**	**100%**	**533**	**100%**

Table 3

Transfer Credits, First-Semester Credits and GPA, Stop-Out Records, and Age at Entry of URM STEM Traditional and Transfer Students

Variable	Student Group	Mean	SD
Number of Transfer Credits	URM transfer	71.8	32.3
	URM traditional	13.1	15.6
Number of Passed First-Semester Credits	URM transfer	11.0	4.6
	URM traditional	13.9	3.7

First-Semester GPA	URM transfer	2.7	0.9
	URM traditional	2.9	0.8
Number of Years with Stop-Out Records	URM transfer	1.6	2.4
	URM traditional	1.7	2.2
Age at entry	URM transfer	24.6	5.6
	URM traditional	18.1	0.5

Note. In this study, 1,187 traditional and 533 transfer students' records (enrolled between Fall 2010 and Fall 2016) were involved in the analysis.

Models

Figures 5 and 6 display hypothesized relationships among predictors (i.e., pre-college credits and college academic achievements), demographic covariates, and outcome variables (i.e., the degree attainment status and the likelihood of retention). The frameworks were informed by the literature review and LS-NSSA's previous research and evaluation results.

Figure 5

Factors Influencing Degree Attainment Status

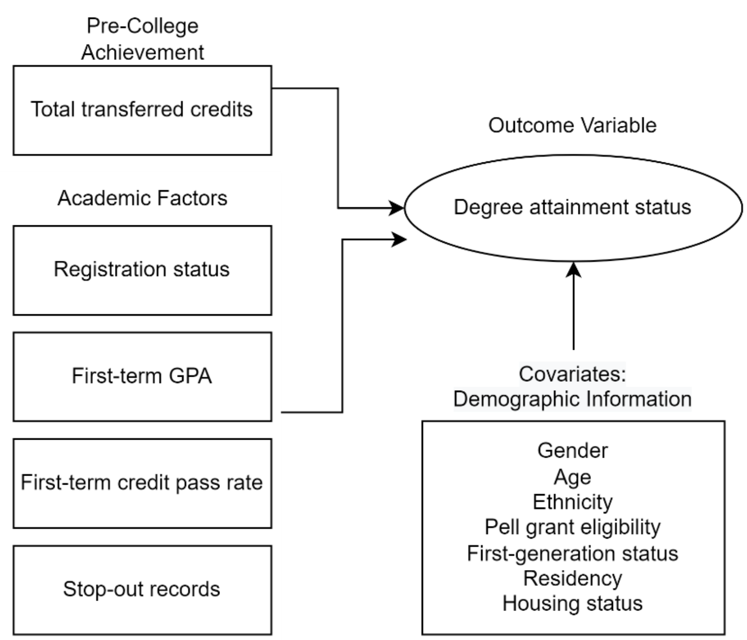
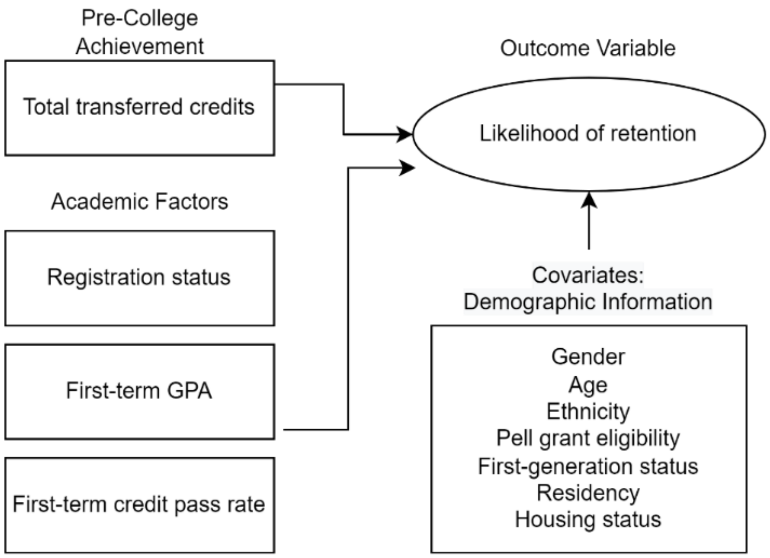


Figure 6

Factors Influencing Likelihood of Retention



Note. These students might still be working toward the degree or might no longer be pursuing a degree.

The hypothesis was that the better URM STEM students adapted to the new environment academically in the first semester, the higher levels of first-semester academic performance they would have, and the more likely they were to be retained in the STEM disciplines even though they were not able to graduate in an expected time frame. Meanwhile, the demographic characteristics, including gender, ethnicity, age, Pell grant eligibility, home location/permanent residency, housing status at entry (off-/on-campus), and first-generation status were included in the models as covariates, controlling for the influence of distinctive socioeconomic backgrounds.

Analysis

Logistic regression methods have been widely used to analyze the association between multiple explanatory variables and educational outcomes, allowing researchers to identify explanatory variables related to interventions and services that contribute to students' outcomes (Peng et al., 2002). Binomial logistic regressions and ordinal logistic regressions were conducted. Binomial logistic models regressed students' degree attainment status—a dichotomous outcome variable—on predictors such as first-semester GPA, first-semester credit-unit pass rate, stop-out records, and other covariates. The ordinal logistic models regressed students' likelihood of retention—an ordinal dependent variable containing four or five values (i.e., zero through three/four years)—on the same predictors, not including stop-out records[2].

Because being both URM and transfer students were assumed to be associated with a greater challenge of retention and degree attainment, for a better understanding of how and to what extent those predictors affected the URM STEM transfer students, the current study differentiated analyses for transfer and traditional students. Four models were tested: (a) Model A—binomial logistic regression of the transfer group's degree attainment status on selected predictors (three-year window); (b) Model B—binomial logistic regression of the traditional group's degree attainment status on selected predictors (four-year window); (c) Model C—ordinal logistic regression of the transfer group's likelihood of retention on selected predictors（three-year window); (d) Model D—ordinal logistic regression of the traditional group's likelihood of retention on selected predictors (four-year window).

The home location variable contained four values and dummy coded into four categorical variables (i.e., the metropolitan area of the state, the greater state area, the

---

[2] The independent variable *Stop-out Records* was an adversely coded variable to the retention records, which was highly linearly correlated with the outcome variable *Likelihood of Retention*. So, it was not included in the ordinal logistic models.

reciprocity states[3], and other states in the United States), indicating the distance from students' permanent home locations to the university. The reference group consisted of students from the metropolitan area of the state. Observations were independent of each other. Variables were approximately normally distributed, though not required in logistic regressions. The full likelihood ratio test was conducted to ensure that the proportional odds assumption was satisfied. Since multiple variables were included in the regression models, multicollinearity analysis was conducted prior to regression analysis. The results suggested no collinearity among predictors (see Appendix C). Absolute correlation indexes were almost all lower than 0.7 (except for the correlation coefficient between the degree attainment status and the stop-out records); the variance inflation factor (VIF) values were all lower than 10; the tolerance values were below 0.2. Therefore, all selected predictors were retained. All statistical tests were conducted via the statistical computing platform R (R Core Team, 2017).

Sub-Study 2 Methods

Sub-Study 2 compared LS-NSSA participants' academic performance (i.e., first-semester GPA, first-semester credit-unit pass rate, and first-year retention) and that of their non-participant and non-URM counterparts.

Design

Retention is associated with interactions among students' educational experiences, psychological perceptions, and sociological outcomes (Hanauer et al., 2016). Findings in previous studies (D'Amico et al., 2014; Dika & D'Amico, 2016) and sub-

[3] Reciprocity states are neighboring states sharing a reciprocal agreement with the state of Minnesota (i.e., Wisconsin, North Dakota, and South Dakota). Based on the agreement, reciprocity students' non-resident admission fees and tuition are reduced or eliminated when attending public colleges and universities in Minnesota.

study 1 indicate that the first-year academic outcomes, especially first-semester GPA and first-semester credit-unit pass rate, have a significant association with URM STEM students' retention and degree attainment. Therefore, sub-study 2 was designed to identify the disparities between LS-NSSA's program participants and their counterparts regarding first-semester academic outcomes and first-year retention rate to address the influence of this program. There were two groups of counterparts—one was URM STEM students who did not participate in LS-NSSA's program and were not financially supported by LS-NSSA; another one was non-URM (i.e., White and Asian American) STEM students who were assumed to be historically privileged in the social and educational system[4]. Table 4 shows the characteristics of these three groups.

Table 4

Characteristics of the Three Groups in Sub-Study 2: Target Population and Their Counterparts

Group	LS-NSSA Participation	URM Student	STEM Student
Group 1—LS-NSSA Participants	√	√	√
Group 2—Non-Participant Counterparts		√	√
Group 3—Non-URM Counterparts			√

Samples

For the integrity and representativeness of the data analysis, the current study still focused on the three partnering campuses of the University of Minnesota. Being the core of the Alliance and serving the majority of participants, the population in this location was representative of the Alliance. The dataset of students' administrative records was

[4] Southeast Asian American students have been supported by other grants at LS-NSSA because they are not recognized as URM students by NSF's definition. They were not included in this study.

the same as used in sub-study 1, which was obtained from the Office of Undergraduate Education, containing students' academic information, demographic information, and retention status. Among 14,180 STEM students enrolled in the selected campuses between fall 2017 and fall 2020, 12,706 (89%) were Whites and Asian Americans, and 1,474 (11%) were URM students. LS-NSSA's participants' records were obtained from LS-NSSA. The dataset contained 80 URM STEM students enrolled between fall 2017 and fall 2020 who received direct financial support from LS-NSSA's first-year scholar program. The first-year scholar program offered a $500 stipend to members in their first year after enrollment with the request for members to participate in program workshops and everts and write reflections. Their program records were matched to the institution-level records using their student identification numbers. The sample size was negatively influenced by the COVID-19 pandemic, which reduced the number of students recruited in the academic year 2020-2021.

Since the full dataset contains 90% non-URM, equal sample sizes were pursued between Group 1 and the other two groups to avoid type I statistical errors and have higher statistical power (Ross & Willson, 2017). Eighty samples were randomly selected from the records of non-participant URM STEM students (N_2= 1,394) and non-URM STEM students (N_3= 12,706) enrolled between fall 2017 and fall 2020, respectively. All these students had full-time registration status and citizenship in the United States. The samples selected in sub-study 2 were different from an exact subsample of sub-study 1. But the students from newer cohorts shared similar demographic characteristics and learning contexts with previous cohorts. The descriptive statistical analysis displayed demographic disparities between program participants (Group 1), non-participant counterparts (Group 2), and non-URM counterparts (Group 3). Table 5 summarizes the demographic information of the samples. Since one of LS-NSSA's goals is to bridge partnering institutions and strengthen the transfer pathways for community college students, Group 1 consisted of a higher percentage of transfer students (36%, compared with 18% and 16% for Group 2 and Group 3). In addition, LS-NSSA supports students from low-income families and historically marginalized families. Therefore, Group 1 contained more individuals who were from Black/African American communities (52%), first-generation students (63%, compared with 41% and 15% for Group 2 and Group 3), and eligible for the Pell grant (64%, compared with 45% and 15% for Group 2 and Group 3).

Table 5

The Demographic Information of Program Participants and Their Counterparts

Demographics	Group 1 ($n_1 = 80$)		Group 2 ($n_1 = 80$)		Group 3 ($n_1 = 80$)	
	Count	**%**	**Count**	**%**	**Count**	**%**
Gender						
Male	41	51%	46	58%	51	64%
Female	39	49%	34	43%	29	36%
Non-URM Ethnicity						
White	N/A	N/A	N/A	N/A	74	93%
Asian American	N/A	N/A	N/A	N/A	6	8%
URM Ethnicity						
Black/African American	52	65%	26	33%	N/A	N/A
American Indian/Alaska Native	4	5%	20	25%	N/A	N/A
Hispanic/Latino	24	30%	34	43%	N/A	N/A
Native Hawaiian/Other Pacific Islander	0	0%	0	0%	N/A	N/A
Age (Range)						
≤ 18	48	60%	57	71%	50	63%
19-24	27	34%	23	29%	27	34%
≥ 25	5	6%	0	0%	3	4%

Registration Status						
New-High-School Student	51	64%	66	83%	67	84%
Transfer Student	29	36%	14	18%	13	16%
Housing Status						
Live On-Campus	35	44%	23	29%	37	46%
Live Off-Campus	45	56%	57	71%	43	54%
Pell Eligibility						
Pell	51	64%	36	45%	12	15%
Non-Pell	29	36%	44	55%	68	85%
First-Generation Status						
First-Generation	50	63%	33	41%	12	15%
Not First-Generation	30	38%	47	59%	68	85%

Analysis

A set of t-tests were conducted using the IBM Statistical Package for the Social Sciences (SPSS, Version 25; IBM Corp, 2017) to determine whether there were statistically significant differences in mean GPA and retention rates across the three groups. Specifically, the Independent Samples t-Test and the Two Proportions Test were conducted to compare the population mean of Group 1 with the means of Group 2 and Group 3. The null hypothesis for these tests was that the population means from each pair of unrelated groups were equal. The Independent Samples t-Test allowed for a comparison of the overall means of each pair regarding students' first-semester academic outcomes (i.e., first-semester GPA and first-semester credit-unit pass rate), and the Two Proportions Test allowed for a comparison of the first-year retention rate across different samples (i.e., LS-NSSA program participants vs. non-participant counterparts, and program participants vs. non-URM counterparts). VanVoorhis and Morgan (2007) suggested that a sample size larger than 30 would be appropriate for the method chosen in the current study.

Sub-Study 3 Methods

Sub-study 1 and sub-study 2 indicated that the early learning experience could be critical for URM students in STEM disciplines. The inquiries and the findings led to the design of sub-study 3. Sub-study 3 was a qualitative study that explored how LS-NSSA's participants perceived the importance of the research mentorship experience, with specific inquiries into the influence of the program on their confidence, interest, belongingness, and commitment in STEM fields, and the counterspace that LS-NSSA created to enhance the sense of diversity, equity, and inclusiveness.

The Conceptual Framework

In LS-NSSA's theory of change (see Figure 4), URM STEM students' decision to stay in STEM fields is associated with their goals, commitment, and interest in STEM learning, which are connected with the levels of students' self-efficacy. *Self-efficacy* is "the belief in one's capability to organize and execute the courses of action required in order producing given attainments" (Bandura, 1977, p. 2), representing one's perceived capability and expectations for success based on prior achievement. This academic term can be interpreted as self-confidence in plain language. LS-NSSA's programs were assumed to improve participants' self-efficacy by contributing to their academic performance, social integration, and a better on-campus climate. This idea is consistent with the Social Cognitive Career Theory (SCCT) framework (Lent et al., 1994). SCCT was developed based on Bandura's (1977) Social Cognitive Theory and has been frequently used when analyzing STEM students' retention and success in STEM disciplines and four-year institutions. SCCT suggests that students' learning experiences, self-efficacy, and outcome expectations are key variables when exploring the development of primary academic interests, educational and career choices, and how academic success is obtained.

SCCT indicates that the integrative model fits students from different cultural backgrounds at different education levels, including traditional and transfer URM STEM students. Lent et al. (2013) found support for the model for students enrolled in Historically Black Colleges and Universities (HBCUs) and Predominantly White Colleges and Universities (PWCUs). Another research assessed a group of engineering students in a Hispanic-serving institution and found the model effective across Latinx students (Flores et al., 2014). Self-efficacy/self-confidence and interest in the STEM major will contribute to URM students' retention in the STEM disciplines and long-term engagement in the STEM fields.

The literature in Chapter II suggests the importance of DEI on campus, financial support, undergraduate research opportunities, and high-quality mentorships for URM STEM students. Within the SCCT framework, contextual factors, such as campus climate and mentorship relationships, may affect a person's career trajectory and may even be the key factors in understanding STEM students' choices. An adverse campus climate creates systemic barriers (e.g., overt and covert racism) that turn into psychological barriers, affecting entrance and retention in the STEM field via the effects on self-efficacy/self-confidence and outcome expectations (Fouad & Santana, 2017). If a counterspace[5] could be available for students who have experienced or are about to experience an unhealthy campus climate, the social discomfort and a sense of not belonging could be reduced

[5] Counterspace is defined as a setting where historically marginalized and underrepresented individuals can promote positive self-concepts and self-development through the "challenging of deficit-oriented dominant cultural narratives and representations concerning these individuals" (Case & Hunter, 2012; Keels, 2020).

(Ong et al., 2018). LS-NSSA aims to create a counterspace for URM STEM students by integrating the critical elements. So, based on SCCT and LS-NSSA's theory of change, sub-study 3 investigated the degree to which URM undergraduates' participation in LS-NSSA's research mentorship programs influenced their self-efficacy/confidence level and outcome expectations, increased their interest in STEM, and strengthened their commitment to STEM education.

The Research Setting—Research Mentorship Programs

During the academic years, LS-NSSA connects URM students with different undergraduate research and mentorship opportunities throughout on-campus networks (e.g., STEM faculty members, student organizations, STEM colleges, and university-wide internship programs). The Virtual Epidemiological (Virtual Ep.) Research Program and Minnesota's Discovery, Research, and InnoVation Economy (MnDRIVE) Program are the two that engage the most significant number of student participants each year, both occurring in the summer (the Virtual Ep. program starts to offer twice per year since 2022, by adding a session in Spring).

The Virtual Ep. program grew out of Tiny Earth Research Internship (Tiny Earth), initiated in 2019 in the Alliance. Tiny Earth originated from the Discovery Institute at the University of Wisconsin—Madison, as a course-based research project. The Alliance adopted the research protocol and performed it in a condensed format outside the classroom. By that time, Tiny Earth offered a part-time paid research internship where URM students could learn research methods, practice lab techniques, and perform research activities related to antibiotic resistance. Participants receive a 925-dollar stipend (contingent on participation and attendance) by growing soil bacteria and studying whether they exhibit antibiotic properties against known bacterial cultures in four weeks. The weekday half-day research schedule allowed students to participate in Tiny Earth and hold a part-time job or enroll in summer-term courses. In the Summer of 2020, as the COVID-19 pandemic outbroke and continued, the lab-based Tiny Earth program had to be moved online. The research protocol was modified to fit the virtual context and renamed to the Virtual Ep. program. In recent years, the online program has focused on the global health crisis. Students are engaged in studying epidemiology and the spread of infectious diseases with two instructors. Instead of conducting hands-on research, students in the Virtual Ep. program are expected to design an epidemiological study, review academic articles, summarize findings, create a poster, and present the results. The research period is extended to five weeks. The two instructors each guide several students and meet with them one on one on a weekly base.

MnDRIVE has been a partnership between the University of Minnesota and the State of Minnesota since 2013. The MnDRIVE projects are the most advanced research projects that serve emerging industries to address grand challenges in each STEM area. Each year, over 30 research projects are funded through this partnership at the University of Minnesota, providing plenty of undergraduate and graduate research opportunities for STEM students. In 2017, LS-NSSA started collaborating with the University of Minnesota Informatics Institute (UMII) and connecting URM students with the MnDRIVE research topics in Robotics, Global Food, Environment, and Brain Conditions. MnDRIVE projects are paid research internships that usually take place in faculty members' labs. In an eight-week research period, each LS-NSSA participant is guided and mentored by one or more faculty members, collaborates with research team members, completes different research tasks, and produces a research poster. They will present the research at the annual NSF Summer Undergraduate Research Exposition (SURE) and other local or national research conferences. Student participants receive a 4,000-dollar stipend funded through financial aid. Students who do not live in the metropolitan area will receive a housing stipend or on-campus housing options.

These Virtual Ep. program and the MnDRIVE program both contain financial support, research, and mentorship components for URM STEM students. They are open for students of all grades and academic backgrounds to apply. The Virtual Ep. program applicants are reviewed and interviewed by LS-NSSA's assistant director to fulfill the program's capacity. MnDRIVE applicants are matched with faculty members based on their research interests. Students meet with matched faculty members prior to program implementation. Because some MnDRIVE projects require higher research skills, preference may be given to students who have completed 45 or more credits by the end of the spring semester. Some faculty members also interview students regarding their needs. In the current study, the author recruited participants from the 2022 summer cohort for the interviews. This cohort was the largest cohort across all program years.

Data Collection

Open-ended, semi-structured interviews have been used to study individuals' experiences, obtain rich and in-depth experiential reflections, collect detailed perspectives from program participants, and produce narrative evidence to examine program effectiveness (J. W. Creswell & J. D. Creswell, 2017; Patton, 2014). The semi-structured interview protocol in sub-study 3 was revised based on the Tiny Earth program evaluation plan developed by the author in 2019, which was used to evaluate the Tiny

Earth program (Summer, 2019) and the Virtual Ep. program (Summer, 2020 - Spring, 2022). Informed by the framework of Critical Race Theory (CRT)—which inspired researchers to recognize and confront racial inequalities permeating the system and influencing underrepresented individuals' success—additional questions were added to address the influence of mentorship and DEI-related concerns. The questionnaire contained four sections (see Appendix D for the full interview protocol):

1) The opening question probed students' expectations and results from the research mentorship programs.

2) The research and mentorship experience section solicited reflections regarding mentor-mentee relationships, program challenges, and mentor support.

3) The diversity, equity, and inclusion section investigated students' understanding of these values and their experiences.

4) The program influence section investigated changes in levels of student confidence, plus their interests, goals, and commitment to STEM education.

Student participants of the 2022 summer Virtual Ep. program and the 2022 MnDRIVE program were all involved in the recruitment process. Students were instructed by program staff that participation in program evaluation activities was recommended but optional. In the kick-off meetings of both programs, the author and LS-NSSA's program coordinators presented the evaluation purpose and process. Recruitment information was distributed via emails during the final week of program activities; two rounds of reminders were sent out in the following two weeks; interviews were scheduled and conducted online through Zoom videoconferencing service (Zoom; Zoom Video Communications Inc., 2022) within 20 days after the end of each program in early July and August 2022. In total, 14 (100%) Virtual Ep. participants and 16 (94%) MnDRIVE participants completed the interviews. Each interview lasted 30 to 40 minutes.

Samples

One MnDRIVE participant rejected the invitation but wrote a short reflection. This reflection was saved but not included in the data analysis, to maintain uniformity in the interview protocol and data collection process across participants. The sample (N = 30) of sub-study 3 was not a subsample of the dataset used in sub-study 1 and sub-study 2 but shared demographic similarities. These students were enrolled in their undergraduate education programs between fall 2016 and fall 2021. Among these 30 interview participants, 15 (50%) were students at the University of Minnesota, and the other half belonged to other LS-NSSA partnering institutions (i.e., eight from four-year institutions and seven from two-year institutions). Six participants of Southeast Asian descent were not recognized as the URM population defined by NSF but made up a large portion of Minnesota's culturally diverse population. About 50% of the overall Asian population, which made up about 5% of the state's population, identified as Southeast Asian, ranking the highest in the nation (the national average is around 21%; Van Dort, 2018). Southeast Asian STEM students, supported by some local funding, were an important cultural group in LS-NSSA. They were all included in the interviews because they shared identities as underrepresented and underserved students pursuing STEM degrees in the higher education system and being LS-NSSA program participants. They contributed to understanding program features and the influence from their culturally diverse perspectives based on personal experiences. Tables 6 and 7 display interview participants' demographic and academic information and the program features that they experienced.

Table 6

The Demographic and Academic Information of Interview Participants

Item	Count	%
Gender		
Male	15	50%
Female	15	50%
Race/Ethnicity		
Black/African American	18	60%
Hispanic/Latino	6	20%
Southeast Asian American	6	20%
Home Institution Type		
Four-Year Institution	23	77%
Two-Year Institution	7	23%
Registration Status		
New-High-School Student	25	83%
Transfer Student	5	17%
First-Generation Status		
First-Generation	16	53%
Not First-Generation	14	47%
Grade		
First Year	6	20%
Second Year	12	40%
Third Year	6	20%
Fourth Year	4	13%
Fifth Year and More	2	7%
STEM Major Direction		
Biological Sciences	18	60%
Health Sciences	4	13%
Engineering	3	10%
Physical Sciences	3	10%
Mathematics	1	3%
Natural Resources and Conservation	1	3%

Table 7

Research and Mentorship Features Experienced by LS-NSSA Program Participants

Program Feature	Count	%
Research Mentorship Program		0%

Virtual Ep. (Virtual, Group-based)	14	47%
MnDRIVE (Lab, Team-based)	16	53%
Research Topic		
Global Health	14	47%
Discoveries for Brain Conditions	9	30%
Environment and Conservation	4	13%
Global Food Ventures	2	7%
Robotics and Advanced Manufacturing	1	3%
Mentor Gender Likeness		0%
Matched	18	60%
Unmatched	12	40%
Mentor Race/Ethnicity Likeness		
Matched	5	17%
Unmatched	25	83%

Analysis

The audio conversations were transcribed into written data through Zoom's voice-to-text service, and the author corrected mistakenly transcribed content manually. The total length of all conversations was around 900 minutes (approximately 15 hours). After de-identifying interviewees, the raw transcripts were uploaded into NVivo (Version 12; QSR International Pty Ltd., 2020). The author read through all transcripts at first to obtain an overall sense of the information. A two-step coding strategy was applied, including a Structural Coding process (Guest et al., 2011) and a Focused Coding process (Charmaz, 2006, 2014; Saldaña, 2021).

Structural Coding is a question-based coding approach framed and driven by specific research topics and inquiries. It is applied for studies involving multiple participants, structured or semi-structured interviews, exploratory investigations, or qualitative hypothesis testing (Saldaña, 2021). During the initial coding step, the data corpus was reviewed by question topics and labeled into words or short phrases—the structural codes. These codes split transcripts into individually coded segments, providing a general picture of the data structure and serving as concepts and sub-concepts that would be summarized into higher-level core categories in the second coding cycle. *Focused Coding* is often used as the second-cycle coding approach that categorically themes the data (Charmaz, 2014; H. Rubin & I. Rubin, 2011). Structured codes were clustered into salient categories and subcategories by their shared common features. Then, subcategories and major categories were organized in a hierarchical structure. Table 8 shows the four primary categories, sub-categories under each primary category, and example structured codes generated from data analysis.

Table 8

Major Categories, Sub-Categories, and Structured Codes (Examples)

Category	Sub-Category	Structured Code
General Experience	Challenges	COVID-19 related issues; unclear structure and arrangement; Tight timeframe; difficult tasks; required; unmatched interest
	Achievements	Learning knowledge; practicing research skills
	Fulfilled expectations	Needs in personal development; being beginner-friendly; building up resume
	Unfulfilled expectations	Lacking hands-on experience; lacking social events
Diversity, Equity, and Inclusion	Definition and value statements	Being accepted and respected; accessing opportunities and resources; having culturally appropriate support; mutual learning
	Perceived campus climate	Positive experience on campus; negative experience on campus
	LS-NSSA being a counterspace	Positive program experience; perceived difference; feeling belonged; having shared identity
Program Influence	Confidence/Self-efficacy in STEM research	Confidence in learning; Confidence in career
	Scientific knowledge and skills	Reading articles; understanding sciences; presenting research findings; communicating in scientific languages
	Benefits in opportunities	Expanding networks; continuing assistantships; being more competitive
	STEM interest	Wishing to learn more; exploring STEM pathways; looking for graduate school information
Mentorship Quality	Mentor-mentee relationship	Mutual learning in the research team; role modeling; being open and flexible; caring; negative encounters

Mentor-mentee likeness	Shared backgrounds; no preference in demographics; professionality matters; sympathy
Mentor support	Instrumental support; emotional support; resource connection

Validity and Accuracy

A validity issue is that different samples were drawn for each phase of the evaluation. It would have been desirable to select the qualitative sample from individuals who participated in the quantitative analysis. However, due to unexpected challenges encountered by LS-NSSA during the pandemic, such as personnel turnover and budgeting issues, the duration of research activities was expanded, and the data collection and analysis were completed over an unexpectedly extended period. While samples were drawn from different cohorts, they shared demographic characteristics, institutional context, challenges, and intervention exposure.

Data rigor and transparency are two criteria of data quality in mixed methods research (O'Cathain, 2010). In the current study, the author sought to address this by implementing each method as instructed, being transparent about the sample sources and sizes, performing measures of validity and reliability, and documenting procedures regarding data collection and analysis. An advantage of a mixed methods design is the capacity to triangulate results via evidence from different types of data and analysis (J. W. Creswell & J. D. Creswell, 2017). The statistical analysis and the thematic analysis produced empirical and experiential evidence that could be utilized to triangulate key elements regarding LS-NSSA's theory of change and investigate factors critical for its target population. However, triangulation was not applied within each analysis phase. The author sought to strengthen the validity and accuracy of the account via peer debriefing. The methods, instruments, and results were reviewed by two doctoral candidates whose research topics were related to culturally appropriate strategies in higher education and measurement of student achievement. These two external reviewers raised questions and suggestions that helped locate the items to be clarified. The dissertation committee's review of the entire study also served as quality control.

In terms of external validity, which relates to the generalizability of the quantitative research and the transferability of the qualitative research, it should be recognized that the current study was performed as an evaluation of the LS-NSSA program. The qualitative analysis diversified the voices though, the study limits the context to a specific program design at a specific geographical location. The author explained the theoretical frameworks of the analysis, provided contextual information, and described the characteristics of samples to clarify the transferability of the results and help audiences determine the applicability of the findings in future practices or inquiries.

Chapter Summary

LS-NSSA's theory of change guided every step in the theory-based evaluation and informed the design of three sub-studies in this explanatory sequential mixed methods research. The first two sub-studies applied regression analysis and t-tests to illustrate the factors that drove URM STEM students' retention and degree attainment and whether LS-NSSA's participants outperformed their counterparts on these factors. These inquiries informed the third sub-study in which LS-NSSA's participants of two summer research mentorship programs were interviewed regarding their perspectives of and experiences in program features that might or might not make a difference to their self-efficacy, interests, and commitment in STEM disciplines. Statistical analysis and thematic analysis were conducted using the data obtained from the institution, the LS-NSSA program, and the students. The following chapter reports the results, findings, and interpretations from each sub-study to answer the three research questions, demonstrating what worked in LS-NSSA's theory of change for whom and how.

CHAPTER IV. RESULTS AND FINDINGS

In this sequential explanatory mixed methods study, quantitative and qualitative data analyses were conducted separately.[6] The factors revealed in the quantitative analysis were triangulated and deepened through the thematic analysis. This Chapter presents the results in the order of the research questions:

Research question 1: What are the factors that influence the retention and graduation of LS-NSSA's target population (URM STEM undergraduate students)?

Research question 2: How do LS-NSSA participants perform on key metrics compared with their non-participant counterparts?

Research question 3: How do participants perceive the importance of the research mentorship experience?

- 3a) How does the experience at LS-NSSA influence their confidence (self-efficacy), interest, belongingness, and commitment in STEM fields?
- 3b) How do participants perceive LS-NSSA's influence on diversity, equity, and inclusion?

Factors Associated with URM STEM Students' Degree Attainment and Retention

Four logistic regression models examined factors influencing the retention and graduation of LS-NSSA's target population (URM STEM students): (a) Model A—binomial logistic regression of the transfer group's degree attainment status on selected predictors (three-year window); (b) Model B—binomial logistic regression of the traditional group's degree attainment status on selected predictors (four-year window); (c) Model C—ordinal logistic regression of the transfer group's likelihood of retention on

[6] To avoid ontological, epistemological, and methodological conflicts between distinct methods (J. W. Creswell & J. D. Creswell, 2017).

selected predictors (three-year window); (d) Model D—ordinal logistic regression of the traditional group's likelihood of retention on selected predictors (four-year window).

Table 9

Four Models Examined in Sub-Study 1

Outcome Variable	Regression Model	URM STEM Student: Transfer Student	URM STEM Student: Traditional Student
Graduation Status	Binary Logistic Regression	Model A	Model B
Likelihood of Retention	Ordinal Logistic Regression	Model C	Model D

Model Statistics

Tables 10 and 11 summarize the coefficient-level and model-level statistics of the four models. Both tables contain the odds ratios to show the effect sizes generated by the logistic regression analyses.

Table 10

Relationship between STEM Students' Likelihood to Graduate and Predictors

Variable	Model A	OR	Model B	OR
Coefficient-Level Statistics				
First-generation Status	* -0.70	0.50	0.00	1.00
	(0.31)		(0.18)	
Male	-0.39	0.69	-0.28	0.76
	(0.29)		(0.16)	
Age	0.01	1.01	-0.14	0.87
	(0.03)		(0.16)	
Total Transfer Credits	*** 0.04	1.04	*** 0.03	1.03
	(0.01)		(0.01)	
First-term Credits Pass Rate	*** 0.03	1.03	0.01	1.01
	(0.01)		(0.01)	
First-term GPA	0.23	1.26	*** 0.82	2.27

	(0.23)		(0.17)	
Greater MN	-0.43	0.65	* -0.52	0.59
	(0.40)		(0.21)	
Reciprocity States	0.63	1.87	-0.19	0.83
	(0.73)		(0.31)	
Other States	-0.24	0.79	0.02	1.02
	(0.50)		(0.26)	
Housing On-campus	0.10	1.10	** 0.55	1.74
	(0.76)		(0.18)	
Pell Grant Eligibility	0.05	1.05	-0.18	0.84
	(0.30)		(0.19)	
Full-time Status	0.89	2.43	—	—
	(0.60)		—	
Stop-out years	*** -5.51	0.00	*** -4.99	0.01
	(1.11)		(0.61)	
Constant	*** -4.96	—	-0.69	—
	(1.26)		(2.98)	
Model-Level Statistics				
N	533		1,187	
Null Deviance	736.32 (532)		1609.25 (1186)	
Residual Deviance	351.87 (519)		895.05 (1174)	
X^2	*** 384.45 (13)		*** 714.2 (12)	
AIC	379.87		921.05	

Statistical significance levels: * p < .05, ** p < .01, *** p < .001

Note. 1) The numbers in the parenthesis are standard errors. 2) OR stands for odds ratio, which is calculated as the exponentiation of the coefficients. 3) Since 99.9% of URM STEM traditional students were registered as full-time students, that variable was not included in Model B. 4) Students having permanent residency in the metropolitan area in this state were in the reference group for home locations in both models.

Table 11

Relationship between Non-Graduated STEM Students' Retention and Predictors

Variables	Model C	OR	Model D	OR
Coefficient-level Statistics				
First-generation Status	0.10	1.11	-0.29	0.74
	(0.27)		(0.16)	
Male	-0.24	0.79	* 0.23	1.26
	(0.25)		(0.15)	
Age	-0.02	0.98	-0.27	0.76
	(0.02)		(0.13)	
Total Transfer Credits	* -0.01	0.99	-0.01	0.99
	(0.00)		(0.01)	
First-term Credits Pass Rate	0.04	1.00	** 0.01	1.01
	(0.46)		(0.00)	
First-term GPA	*** 0.65	1.92	*** 0.84	2.32
	(0.15)		(0.12)	
Greater MN	-0.22	0.80	** -0.48	0.62
	(0.31)		(0.18)	
Reciprocity States	* -1.32	0.27	*** -0.96	0.38
	(0.55)		(0.23)	
Other States	-0.06	0.95	-0.42	0.66
	(0.40)		(0.23)	
Housing On-campus	0.41	1.51	-0.21	0.81
	(0.79)		(0.16)	
Pell Grant Eligibility	-0.25	0.78	-0.06	0.94
	0.27		(0.15)	
Full-time Status	-0.19	0.83	—	—
	(0.45)		—	
Model-level Statistics				
N	285		697	
Residual Deviance	706.53		1872.95	
AIC	736.53		1902.95	

Statistical significance levels: * p < .05, ** p < .01, *** p < .001

Note. 1) The numbers in the parenthesis are standard errors. 2) OR stands for proportional odds ratio, which could be interpreted as for odds ratios in binary logistic regressions. 3) Since 99.9% of URM STEM traditional students were registered as full-time students, that variable was not included in Model D. 4) Students having permanent residency in the metropolitan area in this state were in the reference group for home locations in both models.

The Results of Model A and Model B

On the model level, the Chi-squared test shows the goodness of fit of Model A and Model B: $X_A^2\ (13, N_A = 533) = 384.45, p < .001$ and $X_B^2\ (12, N_B = 1{,}187) = 714.2, p < .001$. In Model A and Model B, students' graduation status in the three-year or four-year window was regressed on a series of academic performance-related predictors by controlling demographic variables. The results indicate that four predictors were significantly associated with URM STEM transfer students' graduation status in the three-year window, and five were significantly associated with URM STEM traditional students' graduation status in the four-year window (see Table 10).

Four variables were significant in Model A. First, the number of years with stop-out records (i.e., the number of times students suspended their studies) was a significant negative predictor. The *more* stop-out records, the higher odds of transfer students being *less* likely to graduate within the three-year window (p < .001), holding all other variables constant. Second, the total number of credits transferred from previous institutions was positively associated with the degree attainment. For every one unit *increase* in transfer student's transfer credits, the odds of being *more* likely to graduate in three years after the transfer was multiplied by 1.04 times (i.e., increased by 4%; p < .001), holding all other variables constant. Third, the percentage of credit units passed in the first semester of enrollment was positively related to the likelihood. One percent higher first-semester credit-unit pass rate increased the odds by three percent ($p < .001$), holding all other variables constant. Fourth, being a first-generation college student in the family was a negative demographic factor. First-generation students were 50% [i.e., (1 - 0.50) * 100%, $p < .05$] *less* likely than non-first-generation students to graduate within three years, holding all other variables constant.

Five variables were significant in Model B. First, the number of years with stop-out records was a significant negative predictor. The *more* stop-out records, the higher odds of traditional students being *less* likely to graduate within the four-year window ($p < .001$), holding all other variables constant. Second, the total number of credits transferred from previous institutions was positively associated with the degree attainment. For every one unit *increase* in traditional student's transfer credits from high schools, the odds of being *more* likely to graduate in four years was multiplied by 1.03 times (i.e., increased by 3%; $p < .001$), holding all other variables constant. The third and fourth variables were demographic predictors related to where they came from (i.e., home location) and where they lived at entry (i.e., housing status). The *closer* the geographical relationship with the campus, the *more* likely URM STEM traditional students would graduate within four years of enrollment. For example, for students who *did* live on campus, the odds of being *more* likely to graduate in time was 1.74 times ($p < .01$) that of students who lived off campus, holding all other variables constant. Regarding residency, students with home locations in the state but *outside* of the metropolitan area were 41% [i.e., (1 - 0.59) * 100%, $p < .05$] *less* likely to obtain the bachelor's degree in time compared with students having home locations within the metropolitan area. The fifth variable was an academic predictor—first-semester GPA. For every one-unit *increase* in a student's first-semester GPA, the odds of being *more* likely for the student to graduate in time was multiplied by 2.27 times ($p < .001$), holding all other variables constant. It had the strongest effect size among all significant variables in Model B.

The Results of Model C and Model D

Model C and Model D were ordinal logistic regression models focused on the students who did not graduate within the three-year or four-year window. Students' likelihood of retention was regressed on the same series of academic performance-related predictors (except stop-out records) and demographic covariates. Of the 533 URM STEM transfer students, a total of 285 students could not graduate within three years after the transfer; 697 out of 1,187 URM STEM traditional students did not graduate within four years of enrollment. The results show that three predictors were significantly associated with transfer students' retention, and five were significantly associated with traditional students' retention (see Table 11).

Three variables were significant in Model C. First, the *higher* first-semester GPA was associated with the *greater* likelihood of retaining in the following academic year. Specifically, for a one-unit *increase* in first-semester GPA, there was a 0.65-unit (OR = 1.92, $p < .001$) *increase* in the likelihood of retention among non-graduated transfer students, holding all other variables constant. Second, the number of transfer credits negatively influenced their retention. For a one-unit *increase* in the number of transfer credits, there would be an 0.01-unit (OR = 0.99, $p < .05$) *decrease* in the likelihood of registering in the following academic year, in the log odds scale, holding all other variables constant. Third, the home location was significantly related to their likelihood of retention. For transfer students from reciprocity states, the odd of registering in the following academic year was 73% *lower* [i.e., (1 -0.27) * 100%, $p < .05$] than students having a home located in the metropolitan area in this state, holding all other variables constant. Other demographic characteristics such as gender, age, living on or off campus, and socioeconomic index (i.e., first-generation status, Pell grant eligibility) made *no* significant difference to their likelihood of retention.

Five variables were significant in Model D. First, first-semester GPA was positively associated with traditional students' likelihood of retention. For a one-unit *increase* in first-semester GPA, there was a 0.84-unit (OR = 2.32, $p < .001$) *increase* in the likelihood of retention among non-graduated traditional students in Model D, in the log odds scale, holding all other variables constant. The second and third variable were related to students' home location. The *further* distance from the home location to the institution, the *less* likely students would retain in the following academic year. Compared with students having a permanent home in the metropolitan area, the odd of retention was 38% *lower* [i.e., (1 -0.62) * 100%, $p < .01$] for students from the areas outside of the metropolitan area in this state, and 62% *lower* [i.e., (1 -0.38) * 100%, $p < .001$] for students from reciprocity states, holding all other variables constant. Fourth, male traditional URM STEM students were 26% (OR = 1.26, $p < .05$) *more* likely to retain than female students. Last, in terms of academic predictors, first-semester credit-unit pass rate was statistically significant for traditional students. For a one-unit *increase* in the credit-unit pass rate, it was expected to have a 1.01-unit (OR = 0.01, $p < .01$) *increase* in the likelihood of retention on the log odds scale, holding all other variables constant.

The Threshold of First-Semester GPAs

To better understand how first-semester GPA was related to URM STEM students' retention, a sub-test was conducted by coding first-semester GPA into four dummy variables based on quartile statistics. For transfer students, the lowest level of GPA ranged from 0 to 2.33; the second to the lowest level ranged from 2.34 to 2.90; the second to the highest level ranged from 2.91 to 3.29; the highest level ranged from 3.30 to 4.00. For traditional students, the lowest level of GPA ranged from 0 to 2.59; the second to the lowest level ranged from 2.60 to 3.06; the second to the highest level ranged from 3.07 to 3.48; the highest level ranged from 3.49 to 4.00. Model E and Model F were ordinal logistic regressions of the likelihood of retention on four dummy variables without adding additional predictors. The result (see Table 12) shows that for transfer and traditional students having the *lowest* levels of GPAs (i.e., lower than 2.33 and 2.59, respectively), the odds of retention was 73% (OR = 0.27, $p < .001$) and 71% (OR = 0.29, $p < .001$) *lower* than students having the highest GPAs. It suggests that 2.33 and 2.59 could be the thresholds to have a higher likelihood of retention for transfer and traditional students, respectively.

Table 12

Relationship between Non-Graduated STEM Students' Retention and First-Semester GPA Levels

Variables	**Model E**	**OR**	**Model F**	**OR**
Coefficient-level Statistics				
First-Semester GPA [LOW]	*** -1.29	0.27	*** -1.25	0.29
	(0.31)		(0.21)	
First-Semester GPA [MID_1]	0.04	1.04	0.04	1.04
	(0.36)		(0.22)	
First-Semester GPA [MID_2]	0.12	1.13	0.28	1.33

	(0.27)	(0.23)
Model-level Statistics		
N	285	697
Residual Deviance	729.57	1961.58
AIC	741.57	1975.58

Statistical significance levels: * p < .05, ** p < .01, *** p < .001

Note. 1) The numbers in the parenthesis are standard errors. 2) OR stands for proportional odds ratio. 3) The groups with the highest first-semester GPA were the reference group.

Conclusion

The first semester of admission for traditional students and the first semester after the transfer for transfer students were critical for their retention and degree attainment. The foundation of the undergraduate study (reflected by the relative credit units earned in previous institutions), the living situations (reflected by the housing status and the home location), and first-semester academic achievement (reflected by the GPA and credit-unit pass rate) had significant implications regarding URM STEM students' decision-making in the following three or four academic years. Retention did not guarantee success regarding degree attainment but less stop-out records predicted the bachelor's degree attainment. I.e., transfer students with inconsistent registration records were less likely to graduate in the three-year window; and traditional students with inconsistent registration records were less likely to graduate in the four-year window, as shown in the analyses.

Comparing LS-NSSA Participants and Their Counterparts on Key Metrics

Group comparisons through t-tests responded to research question 2. Samples engaged in the study were LS-NSSA participants (Group 1) and two counter-groups—the URM STEM students not affiliated with LS-NSSA (Group 2) and non-URM STEM students (Group 3). Group 3 was not the target population of LS-NSSA.

Descriptive Statistics

Descriptive statistics summarize first-semester academic information of the three groups (see Table 13). Figures 7 and 8 display the boxplots that visualized the mean, maximum, minimum, first and third quartiles, and extreme cases of each group's total transfer credits and first-semester GPAs. Apart from the variables mentioned above, the analysis added the first-year retention record (dummy coded as 0 and 1) because it was the first checkpoint of students' retention. LS-NSSA expects participants to achieve a 3.0 GPA—equivalent to a B letter grade on a 4.0 GPA scale, and to avoid a GPA lower than 2.0—equivalent to a C letter grade and lower than the threshold for retention. A 2.0 and lower GPA would make it difficult for URM students to go through the STEM disciplines. Therefore, the table also displays the counts and percentage of students having GPAs above 3.0 or below 2.0. In general, LS-NSSA served more transfer students. Thus, the average number of transfer credits at entry was higher than the other two groups. LS-NSSA participants' performance on all these metrics was better than their non-participant URM counterparts and close to their non-URM STEM counterparts.

Table 13

Transfer Credits, First Semester GPA and Credit Pass Rate, and First Year Retention of Program Participants and Their Counterparts

	Group 1		**Group 2**		**Group 3**	
Academic Information	**Mean**	**Std. Deviation**	**Mean**	**Std. Deviation**	**Mean**	**Std. Deviation**
Average Transfer Credits at Entry	40.77	32.70	26.29	29.89	28.30	25.09
Average First-Semester GPA	3.14	0.62	3.01	0.84	3.31	0.53

	Count	%	Count	%	Count	%
First-Semester GPA ≥ 3.0	52	65%	50	63%	61	76%
First-Semester GPA < 2.0	2	3%	11	14%	1	1%
First-Semester Credit-Unit Pass Rate	N/A	89%	N/A	90%	N/A	95%
First-Year Retention	77	96%	63	79%	76	95%

Figure 7

Boxplot of the Total Number of Transfer Credits by Group

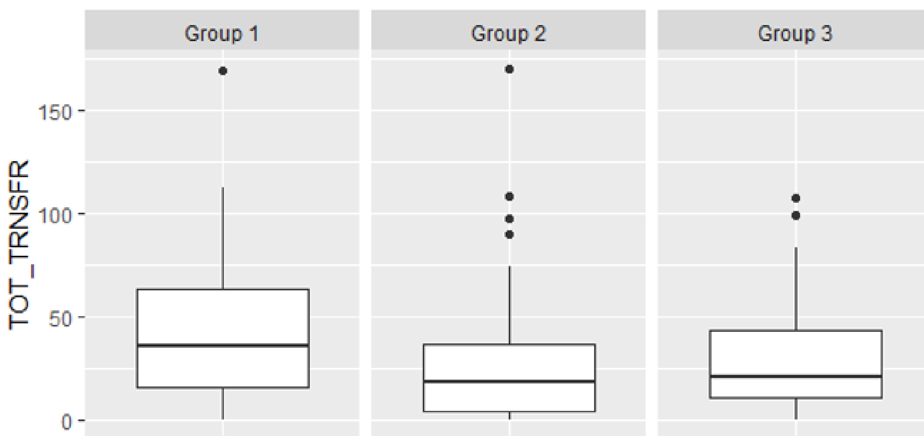
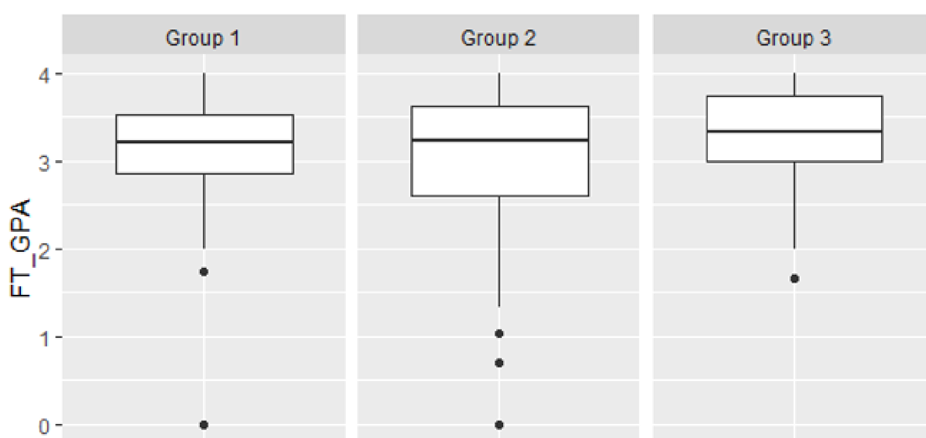


Figure 8

Boxplot of First-Semester GPA by Group



Comparison between Group 1 and Group 2

Descriptive statistics show that, compared with non-participant counterparts (Group 2), program participants had a higher level of average first-semester GPA (3.14 vs. 3.01); a slightly higher percentage of students receiving a first-semester GPA equal to or over 3.0 (65% vs. 63%); a much lower percentage of students obtaining first-semester GPA below 2.0 (3% vs. 14%); a slightly lower level of first-semester credit-unit pass rate (89% vs. 90%); a much higher level of first-year retention rate (96% vs. 79%). The Independent Samples t-Test of program participants and their non-participant counterparts shows that the differences between their mean GPAs and their credit-unit pass rates were not statistically significant (see Table 14). However, the Two Proportions Test suggests that the difference in the first-year retention rate was statistically significant ($p < .001$; see Table 15).

Table 14

T-Tests Comparing Group1 and Group 2 on First-Semester Credit-Unit Pass Rate and First-Semester GPA

				Levene's Test for Equality of Variances		t-Test for Equality of Means			
Academic Outcomes		Mean	Std. Deviation	F	Sig.	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
First-Semester Unit Pass Rate	Program Participants	89.2%	19.2%	0.036	0.850	158	0.851	-0.006	0.303
	Non-participant Counterparts	89.8%	19.2%						
						Mann-Whitney Test			
First-Semester GPA	Program Participants	3.14	0.62	7.611	0.006**	Asymp. Sig. (2-tailed): 0.787			
	Non-participant Counterparts	3.01	0.84						

Statistical significance levels: * p < .05, ** p < .01, *** p < .001

Table 15

T-Test Comparing Group1 and Group 2 on the First-Year Retention Rate

Academic Outcome		Program Participants	Non-Participant Counterparts	2 Proportions Test (Chi-Square)	
				Sig. (Pearson)	Sig. (Fisher's)
First-Year Retention	Yes	77	63	0.001***	0.001***
	No	3	17		

Statistical significance levels: * p < .05, ** p < .01, *** p < .001

Comparison between Group 1 and Group 3

Descriptive statistics show that, compared with non-URM STEM counterparts (Group 3), program participants had a lower average first-semester GPA (3.14 vs. 3.31); a lower percentage of students obtaining first-semester GPA equal to or over 3.0 (65% vs. 76%) and a slightly lower percentage of students receiving first-semester GPA below 2.0 (3% vs. 1%); a lower level of first-semester credit-unit pass rate (89% vs. 95%). However, in the case of being relatively weak in these key indicators, program participants still maintained a slightly higher first-year retention rate (96% vs. 95%). The Independent Samples t-Test of Group 1 and Group 3 shows that there was a statistically significant difference regarding the average first-semester credit-unit pass rate ($p < .05$). There was no statistically significant difference regarding first-semester GPA (see Table 16). Also, according to the Two Proportions Test, the difference in first-year retention between these two groups was not statistically significant (see Table 17).

Table 16

T-Tests Comparing Group1 and Group 3 on First-Semester Credit-Unit Pass Rate and First-Semester GPA

Academic Outcomes		Mean	Std. Deviation	Levene's Test for Equality of Variances: F	Sig.	t-Test for Equality of Means: df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
First-Semester GPA	Program Participants	3.14	0.62	0.495	0.483	158	0.072	-0.164	0.091
	Non-URM Counter-parts	3.31	0.53						
						Mann-Whitney Test			
First-Semester Unit Pass Rate	Program Participants	89.2%	19.2%	7.402	0.007**	Asymp. Sig. (2-tailed): 0.012*			
	Non-URM Counter-parts	94.5%	15.0%						

Statistical significance levels: * p < .05, ** p < .01, *** p < .001

Table 17

T-Test Comparing Group1 and Group 3 on the First-Year Retention Rate

Academic Outcome		Program Participants	Non-URM Counterparts	2 Proportions Test (Chi-Square): Sig. (Pearson)	Sig. (Fisher's)
First-year Retention	Yes	77	76	0.699	0.500
	No	3	4		

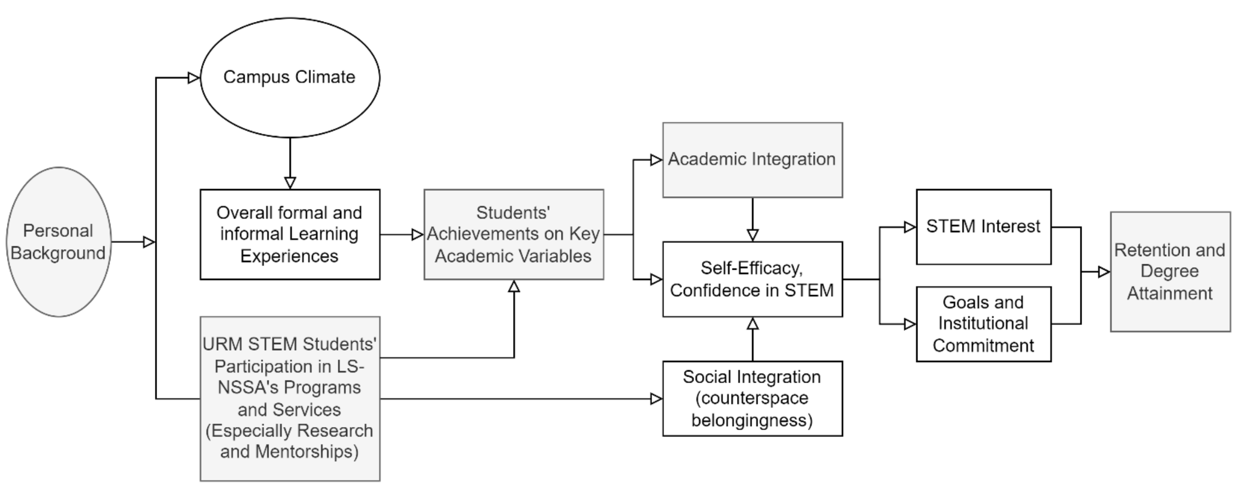
Conclusion

These results suggest that LS-NSSA may have improved URM STEM participants first-year retention rate, contributing to students' consistent enrollment at the beginning of the undergraduate study. Since having a fewer stop-out records was associated with a higher probability of STEM degree attainment for transfer students in the three-year window and for traditional students in the four-year window, it appeared that participation in LS-NSSA program activities may have improved students' academic commitment to pursue STEM bachelor's degrees, compared to URM students not affiliated with the program. Although program participants' mean first-semester GPA was not significantly different from that of their non-participant counterparts, a lower percentage of LS-NSSA participants obtained the GPA lower than 2.0, which has been assumed to be a bottom-line for retention. LS-NSSA participants' academic performance measured in the current study was on par with White and Asian American STEM students. Although the average first-semester GPA and first-semester credit-unit pass rate were lower for program participants, the differences were not all statistically significant. This suggests that observed differences may have occurred by chance ($p > 0.05$).

The results support a portion of LS-NSSA's theory of change (see shaded portion of Figure 9). LS-NSSA program activities, especially the first-year scholar program, appeared to benefit student participants by improving their first-semester achievement and first-year academic integration, which would be expected to translate into improved retention and degree attainment.

Figure 9

Academic Integration Trajectory Evaluated by Quantitative Analysis



The qualitative results suggest that improvements in social integration (i.e., self-efficacy/confidence, belongingness, interest, and goal commitment) may be traced to LS-NSSA program participation.

LS-NSSA Participants' Experience in Research Mentorship Programs

Semi-structured interviews were conducted with 30 LS-NSSA participants in two featured research mentorship programs in the summer of 2022. The results were characterized by four themes:

1. Students' general research experience in research mentorship programs.
2. The quality and influence of mentorship activities.
3. Social-cognitive changes that occurred across student participants.
4. The counterspace created in the research mentorship programs that enhanced the sense of diversity, equity, and inclusion.

Theme 1: Fulfillment and Challenges in the Research Experience

Among the 30 interviewees, 14 participated in the Virtual Epidemiological Research Program, and 16 participated in the MnDRIVE program. The numbers of male and female students were even. More than half came from a Biological Sciences background (60%). First-generation college students counted 53%. Three-fourths were studying at four-year institutions. Five participants were transfer students, and six were Southeast Asian American students. For the Virtual Ep. program, students completed scientific research literature reviews related to each person's interests. The MnDRIVE program sought to match students with projects according to learning interests and focus, in lab settings with varied research tasks.

A Beginner-Friendly Experience. LS-NSSA seeks to comply with NSF grant requirements regarding the demographic characteristics of the target population. Otherwise, LS-NSSA program participation is open to all individuals who fall within the target population, with emphasis on individuals with strong need (i.e., first-generation college students, Pell-eligible students, etc.). Unlike other on-campus research internship programs or undergraduate research assistantships that are usually merit-based, and highly competitive, LS-NSSA's research mentorship programs seek to provide URM students with research opportunities that would otherwise be unavailable. Of these 30 interviewees, 27 (90%) indicated that they had no research experience prior to program participation. This included some second-year community college and advanced university students. Participants appreciated the beginner-friendly accessibility:

> The most valuable experience would be to be in a setting where you are able to reach out to professors and faculty about research and they do not expect you to know everything. I feel that [in other situations] when you go up to professors or ask them stuff, they might expect you to know so many things, but this program was made with the intention that you do not know. So, I appreciate that. (SA, Virtual Ep., July 5, 2022)

Since the spring of 2020, the COVID-19 pandemic has resulted in different levels and aspects of changes in many students' lives. The "stay at home" order, social (physical) distancing policies, and distance learning guidance reduced on-campus research opportunities, making it harder for STEM students to obtain hands-on research experience. Fourteen (47%) interviewees admitted that they were negatively influenced by the pandemic and lacked research skills due to this unexpected event. Under such circumstances, LS-NSSA helped students who intended to be reconnected with the STEM research realm through these entry-level opportunities. As participant EI mentioned:

> Since the year and the year before COVID-19, I did not really have the chance to do anything like research over the summer. So, this summer I wanted to do something different and something where I'll be able to gain from it basically. (EI, MnDRIVE, Zoom interview, August 2, 2022)

The early research experience would be especially beneficial for first-year students enrolled in STEM colleges who have yet to settled in a particular major direction. They would be able to experience research projects as a beginner, explore available paths, and identify their interests via early mentoring experiences:

> I think this was the best position for me, for my first year, because it was my first ever research experience and I really appreciated having a position where I did not need to know a lot or understand a lot of jargon in order to fully participate. (BN, Virtual Ep., Zoom interview, July 6, 2022)
>
> I'm a freshman in college. I really wanted to do something like research or [have] some enrichment opportunity. I applied to a lot of them, but they all ended up rejecting me because they preferred upperclassmen. (SA, Virtual Ep., Zoom interview, July 5, 2022)

Expectations Fulfillment and Challenges. Because the majority of participants were beginners regarding STEM research, their expectations were exploratory: to see how things work in a lab, to learn the process of conducting research—from literature review to data presentation, to connect with faculty and peers and additional resources, and to try out the research-based career path. Most participants acknowledged fulfillment of their expectations. Still, the extent of fulfillment varied due to two significant reasons: (a) Whether or not they were matched with a project that matched their precise research interests or demands, and (b) whether or not they received structured instruction. For example, participant CV (August 6, 2022) mentioned that the work contained a lot of chemistry practice, which did not align with CV's level of math and data science background. Participant FS (August 3, 2022) was initially interested in brain science research, but FS was matched to a global food research team. However, many students still appreciated the value of the research projects that were different from their initial interests:

> I'm a Mechanical Engineering major, and I'm going on to my second part of junior year, so, quiet a way into my college career already. This research was actually about Geology Engineering, very different from what I am studying. But there are things it does help me a ton with problem solving in general. (KHV, MnDRIVE, Zoom interview, August 9, 2022)

The second factor was related to the quality of mentorship. These two part-time summer research mentorship programs had more intensive schedules than semester-long assistantships or full-time summer internship programs. Students were required to complete a study within five and eight weeks and deliver a poster presentation on the last day. Time was too intense to make back-and-forth modifications to the research plan. Although most projects worked smoothly, five (17%) were bothered by unstructured research plans. For example:

> The plan was not complete… I was changing what I was doing those weeks [the first three weeks]. Then, all those weeks passed by, and then you need to get some results. So, in the final weeks [from the fifth week], I was pushed to focus on one thing, which was challenging. I could have just focused on one thing the first week. Maybe that is what research is about. (WA, MnDRIVE, Zoom interview, August 5, 2022)

Except for the challenges brought by the unmatched academic background (n = 5) and the intense research and schedules (n = 6), participants also experienced difficulties due to (a) limited knowledge and skills in research (n = 8); (b) the need to balance between family responsibilities, other job responsibilities, and program activities (n = 4); (c) language barriers (either scientific language or the English language; n = 4); (d) technical and logistical issues brought by external collaborators (e.g., government departments, equipment suppliers; n = 2); and (e) students' physical health problems (n = 1). However, unfulfilled expectations and challenges experienced in programs would only sometimes lead to inferior benefits. It largely depended on the quality of the mentorship, the gains obtained from experience, and the social-emotional needs satisfied through diversity, equity, and inclusion, as the following parts explained.

Theme 2: Relationships, Similarities, and Interactions Defined the Quality of Mentorship

The mentorships in the Virtual Ep. program and the MnDRIVE program were different. In the Virtual Ep. program, students were assigned into two groups, and one instructor instructed each. The instructor was recognized as a mentor for the group members because the group size was small (seven per group), and there were weekly one-on-one meetings between the instructor and each student. The closeness between the Virtual Ep. program instructors and their students was strong. Therefore, students' mentorship experience was not analyzed separately by program. The results show that the helpfulness of the mentorship was not defined by the program design. The genuine care from the mentor, a structured mentoring style, and the combination of instrumental and emotional support from the mentor mattered, especially when challenges occurred.

The Mentor-Mentee Relationship. Among all participants, two explicitly expressed their dissatisfaction with their mentorships—one for the sense of alienation, the other for the mentor's unpreparedness to deal with cultural differences. Participant EM (August 3, 2022) described how busy the mentor was. During the eight-week experience, the mentor could not deliver support effectively and did not assign a co-mentor (usually a graduate assistant of a mentor) to help with EM's project. EM hesitated to reach out to the mentor because "[the mentor] was going to be very busy" (EM, August 3, 2022). This situation generated a sense of alienation and isolation. KT (August 3, 2022) also said that "I did not really talk to [my mentor] that much because I knew [my mentor] was busy" (KT, August 3, 2022). Insufficient communication generated additional issues when there was a culture shock for both the mentor and the mentee, as perceived by KT. Being the first MnDRIVE mentee on the team, "[the research team] was not quite ready for me, and they did not have things quite set up the right way," said KT (August 3, 2022). There was a longer mutual learning process than other participants experienced because of a cultural difference between KT, the mentor, and the research team. The different cultural and educational backgrounds generated communication issues—partly due to the English language, partly due to different communication styles, and partly due to the unmatched expectations. KT felt it difficult to "succeed in *[my mentor's] eyes*, as well as *my eyes*" because "there was a lot of pressure on the individual to have everything together and know everything, instead of being able to rely on the group" (KT, August 3, 2022).

The Mentor-Mentee Similarities. KT's experience indicates that different cultural backgrounds between a mentor and a mentee could influence the effectiveness and closeness of the mentorship. However, this argument did not receive additional support from other interviewees' experiences. About 18 (60%) participants shared their gender identities (i.e., male or female) with their mentors, and five (17%) shared their ethnic and cultural backgrounds with their mentors. When asked whether gender and ethnic disparities negatively influenced the mentor-mentee relationship, neither did participants who shared demographic similarities with mentors nor those who did not agree on it. For mentees, a mentor's professionality in the research area mattered, and the importance of the similarity in research interests exceeded the demographic similarities. Some participants even saw demographic disparities as a benefit:

> I think [sharing cultural background] would help with the comfortability, but at the end of the day, I would say to me that does not really matter, because once I really go into the workforce or the career path, I will be meeting people with different backgrounds, and I should start getting used to interacting with people outside of my culture and outside of my comfort zone. (CV, MnDRIVE, Zoom interview, August 6, 2022)

The demographic disparity was not on the list when talking about the expected characteristic that a future mentorship would better have. As participant EC (July 13, 2022) indicated, mentees cared about a mentor's "knowledgeable about the things that [mentees] would like to do and to study in the future." A mentor's passion for the research study, curiosity about a research problem, and excitement in solving a problem could inspire the sense of "similarity" with mentees. As KZ (August 9, 2022) shared: "I shared that excitement when I got those results and [the mentor] was also very interested in it. That was something that we could with easy for us to connect on."

Also, sharing research interests and career goals could help picture a role model for mentees to have a preview of their chosen profession— "A mentor is at a point in your career where you want to be in the future, and you will learn from them," said NA (August 4, 2022). In these research mentorship programs, not only faculty mentors became role models for participants but also their graduate assistants. Usually, when faculty mentors were busy, the graduate assistants provided instructions, answered questions, helped with technical issues, and shared resources with undergraduate mentees. They understood the challenges a student researcher might encounter, and they would share tips based on personal experiences. Some participants might even have MnDRIVE alumni as co-mentors:

> [One of the co-mentors] was actually part of MnDRIVE last summer, who completed the research project, and we did meet a few times that talk about the experiment that [the co-mentor] did, which was very helpful and so I had an additional contact person. (SH, MnDRIVE, Zoom interview, August 10, 2022)

This co-mentorship, or peer mentorship, added another layer to the sense of similarity between mentors and mentees.

The Mentor-Mentee Interactions. Regarding the qualities that mentees appreciated in the mentorship, the most frequently mentioned were mentors being “supportive” (73%), “flexible” (67%), “responsible/approachable” (53%), and “open” (33%). These qualities were highly related to the meeting frequency between mentors and mentees. Participants EM and KT, as mentioned above, felt isolated and insufficiently supported because of their mentors’ unavailability. On the contrary, 22 participants who were satisfied with their mentors’ support and 16 who appreciated their mentors’ being responsible and approachable usually had more frequent meetings with mentors (i.e., once or more than once a week). The meeting format was usually diverse (i.e., individual sessions, small group meetings, and formal lab meetings). For example, participants such as CU and CHV kept continual connections with their mentors:

> We would have meetings every once in a while, when we had something to share, if not at a formal lab meeting. We would schedule independent meetings with just a couple of us, and we would talk about the things that we want to do or things that we want to try. Those were super helpful. (CU, MnDRIVE, Zoom interview, August 9, 2022)
>
> [My mentor] was always available whenever I needed help or whenever I had questions. I know that [my mentor] had a very busy schedule but would always made time for me. Almost we met two times a week… [My mentor] described the busy schedule to me yet [my mentor] still had a lot of time to help me out. (CHV, MnDRIVE, Zoom interview, August 5, 2022)

Participants having positive mentor-mentee interactions described more positive learning experiences and personal development plans than EM and KT did. Table 18 presented some examples of participants’ experiences regarding positive mentor-mentee interactions, by mentorship quality types.

Table 18

Qualities that Mentees Appreciated in the Mentorship in Zoom Interviews

Mentorship Quality	Example Quotation
Being Supportive (i.e., deliver instructions and resources)	My mentor was really good about getting one-on-one time outside of class. if we needed help and then [the mentor] would come around towards your rooms and asked [how to help]. (ML, Virtual Ep., July 7, 2022) I feel very comfortable going to [my mentor] and communicate. I would say we have a good communication. We make sure to set time aside for us to check up, go through tough questions, and explain things. (SZ, MnDRIVE, August 2, 2022)
Being Flexible (i.e., with schedules, arrangements, research content, etc.)	I got so many other things going on, personal and medical, that I always get the times confused. But [my mentor] was gracious enough to accommodate me and allow us to meet anyway when [my mentor] was able to do that. (BN, Virtual Ep., July 6, 2022) [My mentor] was able to accommodate to what I was interested in, and the lab was able to offer things like that. So, [my mentor] switched [the setting] around to be more of a biological lab component for me, and it worked perfectly. I think I just got lucky. (FS, MnDRIVE, August 3, 2022) I also worked part time outside of this. So, there was flexibility according to my own schedule and there was time throughout the week to set up meetings and for me to come into labs and go over things. I do feel lucky. (RZ, MnDRIVE, August 5, 2022)
Being Responsible/ Approachable (i.e., to meet, talk, listen, and help)	[My mentor] does not push me off to somebody else. I can go directly to [my mentor] and [my mentor] is more than happy to help and especially if there is something that is not necessarily what I am focusing on with my project. (KZ, MnDRIVE, August 3, 2022) The instructor did a really great job at keeping touch with every single person and it feels that [then instructor] more of a teacher or an instructor. [The instructor] was for sure a one-on-one mentor. (RC, Virtual Ep., July 6, 2022)

Being Open (i.e., to ideas, opinions, explorations, and interests)	My mentors asked me what I wanted to do. I did have a list of options, and I could choose any of them that I felt comfortable with. (SH, MnDRIVE, August 10, 2022) There were different things you could have done in the lab. So, we do a lot of experiments and behavioral tasks, but then there were other aspects like coding or building stuff, knowing how different instruments work. There was imaging data and there was a lot of different parts of what we do in the lab. So, I was just able to try those things. (SZ, MnDRIVE, August 2, 2022) Anytime when I came up with an idea, [my mentor] would always be willing to try it out. Sometimes [my mentor] knows it would not work, but [my mentor] just wants me to get some experience. So, if I come up with an idea, [my mentor] lets me do it anyways without saying anything. (KHV, MnDRIVE, August 9, 2022)

As quoted from participants in Table 18, formal interactions on research projects and informal interactions outside the classroom or the lab developed positive mentor-mentee relationships. Sometimes, a positive informal interaction would promote formal interactions. For example, knowing about "[team members'] kids and their pets and the things they like" made MB (August 11, 2022) more comfortable to reach out to the mentor and co-mentor and ask questions about work and life; the "small community" built in the lab brought the sense of trust towards the mentorship to SZ (August 2, 2022); apart from the communication on research projects, WA also received advice in "general academic and professional development issues" from the mentor (August 5, 2022), which strengthened the mutual understanding.

Support Obtained from Mentors. Previous research has examined different types of faculty mentorship, including the instrumental mentoring approach (i.e., with a focus on knowledge and skills development, improving mentee's competencies, and supporting mentees to pursue and achieve academic goals) and developmental/emotional mentoring approach (i.e., with the focus on the emotional bond between mentors and mentees by spending time together and increasing closeness; Eby et al., 2007; Schenk et al., 2020). The current study also investigates the types of support mentees received from LS-NSSA mentors.

Instrumental Support. Within the five and eight weeks of these research mentorship programs, primarily, mentors delivered project-based instrumental support, including: helping mentees select and shape research topics; teaching mentees about research methods and tools; guiding mentees through the research process; providing academic and technical support required to complete the projects; suggesting poster development and presentation. All participants reflected that their knowledge and skills in STEM research improved—from different aspects and to a different extent. It depended on what research tasks they completed with the research team.

For Virtual Ep. participants, their capabilities in raising research questions, reviewing research literature, summarizing research arguments, and presenting results improved. They learned to use scientific language to communicate epidemiological research projects. Since the topics were decided by participants—with suggestions from mentors—based on their study and career interests, the knowledge and skills obtained from experience directly benefited their academic goals and career path development. For MnDRIVE participants, their development in STEM research-related knowledge and skills was multidimensional, from literature review to experiment design. Their experiences were more hands-on and contextualized, with more training in troubleshooting and problem-solving skills. Structured research plans contributed to more effective instrumental support, perceived by participants.

Developmental/Emotional Support. Considering that participants lacked research experience and had diverse challenges during the programs, there was a need for developmental and emotional support. Sometimes, a word of encouragement could comfort participants in a stressful situation. Nine mentors were available emotionally, perceived by participants. As some participants shared: "[My mentor] was always telling me how good I was doing, and how well the project is going to come, and we are going to complete it on time" (BM, August 4, 2022); "[My mentor] was really proud of what I did" (KZ, August 3, 2022); "they made me feel that my ideas were very well valued" (KHV, August 9, 2022); "they would say, 'do not worry about it and we can just help you out with this'" (NA, August 4, 2022). Though experiencing cultural shock in the lab, KT appreciated the emotional effort that the mentor made to ease the tension:

> One day we got really upset and really frustrated, but then the next day, we came in and [my mentor] started really saying, "Good job. Good job" about a lot of stuff, which I felt like [my mentor] was restating the position and finding a place and bringing down the frustration, just so we can get to a better spot and so I felt that the relationships were slowly built. (KT, MnDRIVE, Zoom interview, August 3, 2022)

Resource connection set between instrumental support and developmental support. On the one hand, it allowed mentees to upgrade their toolkits and expand their networks in the STEM fields when approaching their goals, while on the other hand, it strengthened the bond between mentors and mentees through a sharing and learning process. Resource connection was built upon a mutual understanding and based on caring, which was out of a research mentor's obligation. However, the two Virtual Ep.

mentors shared campus-wide academic resources (e.g., library access, faculty members, administrative resources) with participants. They offered to connect participants with anything they would request in the future. Five MnDRIVE mentors actively connected mentees with other faculty members, online educational resources, and ongoing research projects of the mentee's interest. Participants FS' and KZ's experiences serve as examples:

> [My mentor] would literally just stop by the lab and say, "Hey! I am going down to donuts and coffee. You should come with me, and I will introduce you to some people," and really giving me the networking opportunity, and I would just be talking to those people, and they would ask, "Hey, what are you interested in?" "What do you want to do?" … It was not quite uncommon, but I definitely got an introduction to many people around the labs… [My mentor] really built that connection and trust. (FS, MnDRIVE, Zoom interview, August 3, 2022)
>
> [My mentor] was more than happy to have me come and take a look at other people's projects in the labs and what they were working on, to get familiar with equipment, for example, using the specific microscopes we have. So, that was fun. [My mentor] also introduced me to other labs. There was a larger lab setting with [Professor C's] lab as well as [Professor B's] lab, and there was a bunch of other different professors, and I got to sit down and watch other labs presenting what they were doing. That was very helpful. [KZ, MnDRIVE, Zoom interview, August 3, 2022]

Theme 3: Changes in Participants' Confidence, Interest, and Commitment in STEM

As a result of the positive experience in research and mentorship, participants reflected that their confidence, interest, and commitment to STEM education and career increased, which were critical elements connecting URM STEM students' participation with their retention and degree attainment.

Improved Knowledge and Skills-Based Confidence. Twenty-seven (90%) participants described their improvement in research knowledge and skills as the top gain from the research mentorship programs. For entry-level student participants, their self-reflected improvement lay in micro aspects, such as understanding what research is, searching for and filtering research topic-related articles, getting more familiar with scientific writing rules, creating PivotTables to analyze worksheet data, putting a presentation PowerPoint together, becoming more skillful in using microscopes, and managing time in a lab-setting. Upper-class participants obtained advanced skills, such as starting and concluding an experiment, using appropriate analytical programs (e.g., JMP Statistical Software) and codes to deal with clinical data, making complex graphs to present research results, getting a study design approved by the institutional review boards, and handling techniques such as PCR (i.e., polymerase chain reaction for DNA analysis).

"I felt that I was not as good because I did not really have the experience" (BM, August 4, 2022)—a bunch of participants felt self-questioned as BM did beforehand. While after the program, participants' minds changed to that— "I was really good at working there!" (KZ, August 3, 2022) In the end, 13 out of 14 Virtual Ep. participants indicated that they become more confident in coming up with research questions, communicating a research idea to an audience, creating a research poster, and reviewing scientific literature. Sixteen out of 17 MnDRIVE participants gained confidence in dealing with challenges in STEM education and employment, collaborating with STEM scholars and practitioners, recognizing their talent in STEM, and so on.

Participants became confident by learning to appreciate and admire themselves. RZ (August 5, 2022) realized that everyone got their own pace in learning. Thus "not everything is going to come out perfectly; not everything is going to come as everyone wants." Therefore, it became essential to recognize what they have achieved, appreciate the efforts taken towards it, and step out for another triumph. As SY (July 7, 2022) said, "I researched something I was really curious about, and I was really proud of what I was able to create at the end and that definitely boosted my confidence." Self-appreciation was particularly important for participants who experienced an untraditional pathway to their current place. NG (July 5, 2022) was one of them:

> I watched my presentation video over and I am like, "Wow! You did a good job! You stay straight to the facts!" Because Lord knows, I wanted to go on a tangent… The data that I got, I built my own brand and started throwing out my own numbers and I was like, "Girl, you could really do this!" … My confidence went up 100 times more than when I first walked in because I did not know what to expect. (NG, Virtual Ep., Zoom interview, July 5, 2022)

Increased Interest in STEM Exploration. No matter which grade participants were in, these two programs invited those who were "outside the room" to get "inside the room" (BN, July 6, 2022).

In the Virtual Ep. program, 13 out of 14 interviewees demonstrated their increased interests in epidemiological research and other learning opportunities: NG (July

5, 2022) described the literature review process as "a train of thought" that stirred up the interest in exploring more ideas and writing up a research proposal; JS (July 5, 2022), a transfer student who was about to start the university life soon, would like to "take some science courses in the first semester"; MM (July 15, 2022), preparing for the transfer, felt prepared to apply for other research programs, such as MnDRIVE, in the university; NT (July 5, 2022) started thinking about transferring because only four-year institutions could provide what NT wanted to focus on—the Environmental Science. The program exposed ZI (July 5, 2022) to the scientific work that many people had been working on, which made ZI realize that "how [scientific research] affects people's long-term life."

MnDRIVE connected participants with more near-future opportunities that matched their research interests. "Once this summer was coming to an end, [my mentor] talked about the chance of having me continue with them," said EI (August 2, 2022). The contract extension went to six other MnDRIVE participants. For participants who were not going to continue the work in the previous labs, they also benefited from connections and resources from the program. Six of them indicated that they had a plan for additional explorations. For example, FS (August 3, 2022) had an interest in Medicine, and through connections and opportunities fostered in MnDRIVE, FS would start another research project at the Mayo Clinic. Such continuous affiliation with research opportunities reinforced participants' sense of worth and turned an exploratory interest into a commitment to STEM.

Strengthened Commitment to Degree Attainment. Among 30 interviewees, 16 (53%) firmly expressed that they were more determined to study and work in STEM. Another nine (30%) had a goal of pursuing a pre-medical program or a medical degree—

which contained STEM-related knowledge, skills, and experiences but were not counted as STEM disciplines for now. Another five (17%) participants, either had no solid plan in their first year or planned on a career with an associate degree in a two-year partnering institution.

The 16 STEM students showed commitment to STEM fields. These programs showed the pathways in specific areas, envisioned a higher degree in STEM, and inspired them to pursue a more significant influence in society. BM (August 4, 2022) felt "meant to be in STEM" and "meant to pursue something there"; SZ (August 2, 2022) confirmed that the research mentorship experience had a "huge influence" on the decision to stay in the science field and solidified what could do; FO (July 8, 2022) aimed at working in the health field, and the Virtual Ep. program helped narrow the focus to "something as an epidemiologist" and to study on diseases; After hearing about the life in graduate programs and working with graduate students in the labs, CU (August 9, 2022) added graduate school into the post-graduation planning and before that, a bachelor's degree would be a necessity; with a STEM degree, "you would advance yourself, and then to impact the community around you" (WA, August 5, 2022).

For some participants, a commitment to degree attainment might not lead to a commitment to a STEM career path. EO (July 6, 2022) was "definitely motivated and driven to get a bachelor's degree," but in terms of careers, EO would "keep the horizon broad." However, early exposure to STEM research would help first-year URM students to "get a head start" and "make it easier to decide on a major" and then "think about the career after that" (EO, July 6, 2022).

Theme 4: Diversity, Equity, and Inclusion—Counterspace Created by LS-NSSA

Recognizing disparities and inequities, discussing them, and not avoiding discussing them is the core advocacy of applying the Critical Race Theory (CRT) lens in research (Ladson-Billings, 1998). Serving historically underrepresented populations and aiming at a greater scope of change on and off campus, LS-NSSA has been addressing inequity issues in its efforts, including the research mentorship programs. BN (July 6, 2022) mentioned a scenario when a guest speaker was invited to present. The speaker shared personal experiences and showed transparency as many others did:

> [The speaker] talked about the mental illness challenges, how he grew up in a very privileged community in the New England area… those type of things really was impressive to me because you do not really have to share those things, not disclose who you really are. But their ability to disclose was really helpful because it is nothing that you do without sacrificing something if you want to go into research. (BN, Virtual Ep., Zoom interview, July 6, 2022)

Participants' Understanding and Expectation of DEI. The interviews inquired about the interviewees' understanding of diversity, equity, and inclusion from their experiences and explored what they valued the most when discussing DEI—considering it as a whole concept. "Having equal access to opportunities and resources" ranked as the first aspect that popped out of participants' minds in interviews (see Figure 10), indicating that elements that promoted academic success and professional development were mainly cared about by students of all backgrounds. Especially for URM students, "when you did not get resources, it would be so difficult to succeed" (KT, August 3, 2022).

Regarding "equal access," some of the participants described it as "making sure that everyone was on the same playing field" (EO, July 6, 2022) because they believed that "anything one person could do, another person could also do equally as well" (KZ, August 3, 2022). But more importantly, there should be additional consideration of the environment, allowing resources to lean towards the ones who were underserved historically—to "give a chance to people that were less deserving" (BM, August 4, 2022), "helping those who were a minority to reach those positions" (SZ, August 2, 2022), and "providing everyone with what they needed" (SH, August 10, 2022). Because being historically underserved and looked down upon, URM students had been "trying to keep up with everything" (MB, August 11, 2022); thus, they "needed different things to reach the same place" (SH, August 10, 2022).

Figure 10

Frequency of the Aspects that Interviewees Addressed Regarding DEI

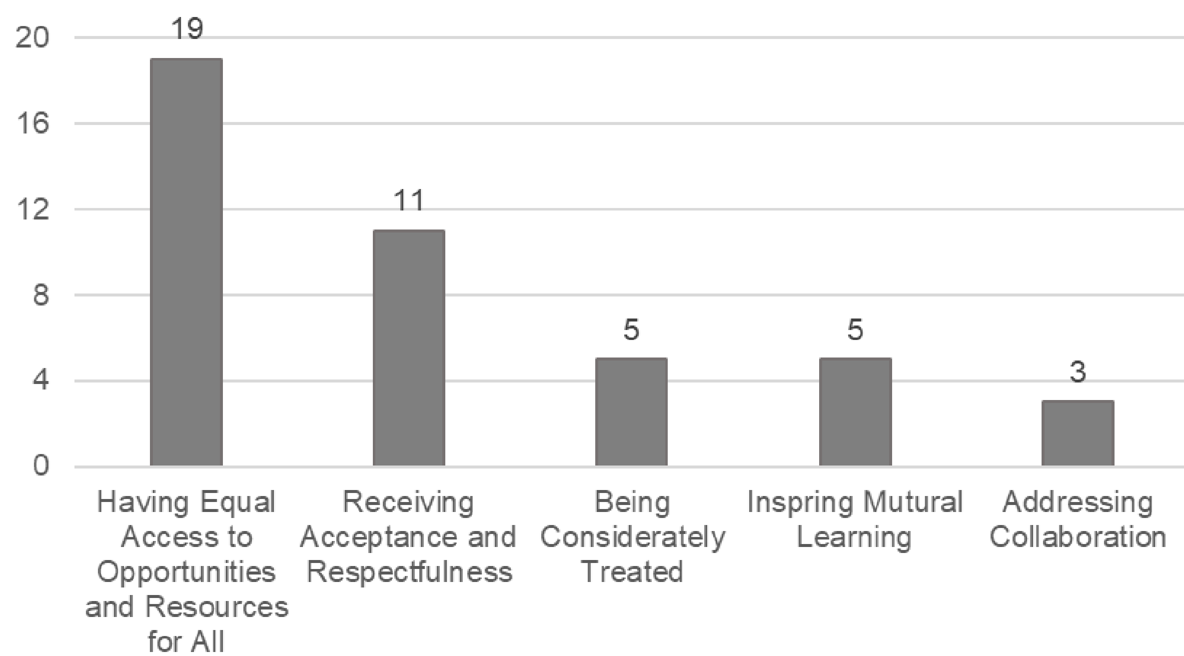


Figure 10 displays other aspects regarding participants' understanding of diversity and inclusion. "Receiving acceptance and respectfulness" sets the tone for the rest. Some participants pointed out that acceptance should be generated from an understanding that people "were all different, in a way" (SH, August 10, 2022)—in some ways, things could change; in other ways, things such as identity, religion, and race could not change. Therefore, acceptance was not about the agreement but respecting the differences. Moreover, participants believed that being diverse was more than accepting differences. It was important for URM students to "see someone of the same color and background working and learning in the same field" (FO, July 8, 2022), but it was equally important to see "not just people who looked like us but people who looked different" (KT, August 3, 2022). A healthy campus climate should consist of these two layers of acceptance.

However, the current campus climate did not live up to that expectation. SA (July 5, 2022) talked about what was observed and what should be achieved from a culturally diverse perspective:

> Diversity comes from having the people that you see outside to be inside. If you were to go outside in the community, the people that you see out there should be reflected where you go in higher education. [However], if you look like the majority of their students, they do not reflect the [city's demographics]. They are completely different. It is because higher education is so unattainable for certain groups. So, I feel that diversity is trying to reflect the best that you can and represent what you see outside. (SA, Virtual Ep., Zoom interview, July 5, 2022)

Negative Campus Climate Experienced by URM STEM Students. About half of the participants (n = 13) described negative experiences that they had experienced on campus in the forms of microaggressions, disproportionate representation, limited exposure to resources, and being forced to be self-dependent. Their lived experiences were better shaped through their language than the author could. The stories are presented in Table 19.

Table 19

Adverse Campus Climate Experienced by URM STEM Students

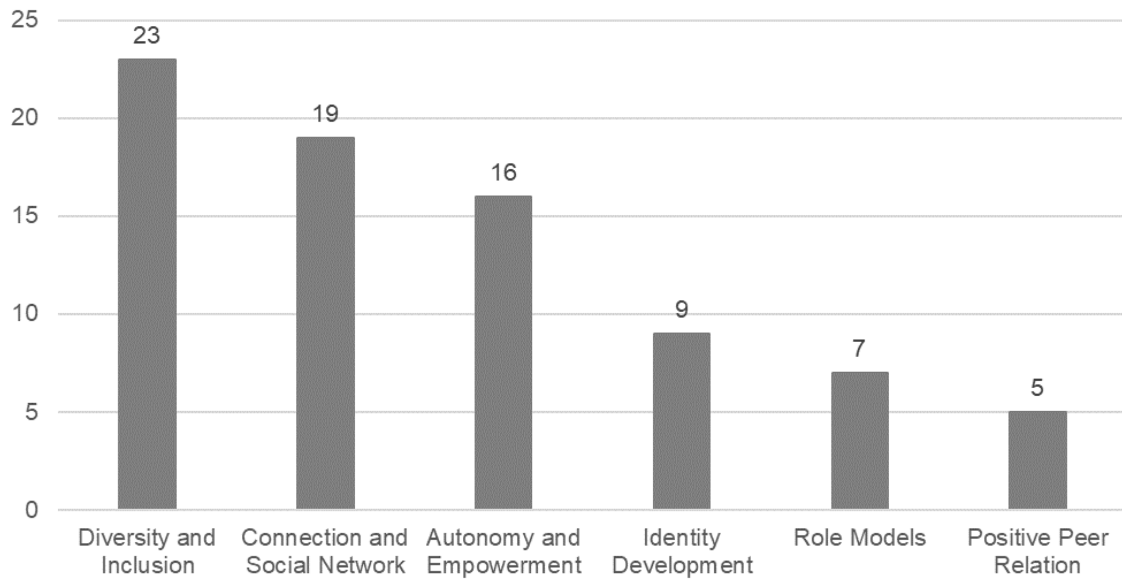
Type of Issue	Example Lived Experiences
Microaggressions	I took lab for chemistry classes, and sometimes I felt that I was being undermined. Just because of the way I look, my knowledge or my capability was being underestimated, for any reason. I also talked to people who are, like me, a Black Muslim girl, who were in those chemistry classes, and they said they felt the same thing too. So, it was a collective experience. (EO, July 6, 2022) some people will say things like, "Oh, you must have had such a hard upbringing. It is so powerful where you are today!" But I think that they do not understand that there is a lot of dualities in people just because you look a certain way, does not mean you experience a certain thing… Sometimes, just because I am a female, and I am Black, and I am Muslim, that I have had a hard life, but I am grateful for the life I had. (SA, July 5, 2022)
Disproportionate Representation	Sometimes I do feel a bit awkward in classes. In my major, it is very dominated by White males. So, me being one of the only girls, not White, sometimes it does make you feel a bit uncomfortable and a bit awkward. (EC, July 13, 2022) I do go to a predominantly White school, so it was really hard maneuvering around it for the last four years. I have been in that institution. So, there was multiple occasions where things were just very uncomfortable for me or I just did not feel like I wanted to continue, especially in the biology field in that school because most of the people in that field are predominantly White individuals. So, I didn't really see many people me or. And it was harder for me to work with people that did not feel related to me. For a lot of people of color, they do not really know exactly where to go to get the [research program] information. (FO, July 8, 2022)

Limited Exposure to Resources	Sometimes I wish there was more for people to do during the summer or even after the summer, during the school year. (EI, August 2, 2022) It just makes it especially difficult. Financial resources are very difficult to find. Opportunities are pretty difficult to find, especially Minnesota is very cliquey. (KT, August 3, 2022) I am new to the country—well, I have been here almost about four years and not having the knowledge of where is where, and not being able to go to high school here and there… All the resources that are available, you have to look for them in a sense. They are not like, "Hey, the school offers this!" There are no outreach programs to show you [the resources]. You have to go look around and find those resources. (NA, August 4, 2022)
Being Forced to Be Self-Dependent	When you grew up in a certain family and all the women in the family are strong, nobody is around here talking about their feelings. They do not have time. My mom did not have time to do that. She was raising two kids by herself. My grandma was the same where she basically raised her kids by herself, too. Talking about feelings and making sure everything is okay—it is just not something I was raised doing. So, I just deal with things on my own because everybody got stuff going on, and I cannot depend on someone to pick up my problems when they are going through there… If I want to be a part of something, I put myself in a position. I am not depending on anyone to give it to me or make room for me. It is just not a conversation I have in my head. I show up to school and even if I am the only Black person in my class. (NG, July 5, 2022)

Belongingness Generated from the Counterspace. Participants perceived that LS-NSSA's research mentorship programs created a counterspace for URM STEM students where they felt safe with people of rich cultural backgrounds, felt accepted and respected by faculty members and peers, developed stronger connections with mentors and expanded social networks, obtained freedom and support to explore what they were interested in, felt empowered to express their ideas and opinions, built academic confidence by viewing the success of similar others, and generated belongingness and science identity in STEM fields (see Figure 11).

Figure 11

Frequency of the Elements in the Counterspace that Fostered URM STEM Students' Belongingness



Trusting relationships with the research teams strengthened participants' science identity and belongingness. They were not treated as an outsider but engaged, trusted, and respected as anyone else who should be with STEM competencies. Such an "empowering" environment stimulated participants' sense of membership/ownership and granted them stronger self-confidence in STEM education:

> They [the mentor and co-mentors] were really taking me in their meeting streaming like *I had been one of them*, and I was working alongside them that I had done graduate research also… I was never looked down upon. (FS, MnDRIVE, Zoom interview, August 3, 2022)
>
> They merely gave me the idea, and then I ran with any idea, and I made an experiment out of it, and they supported me through the whole way. They gave the input; they gave me advice; they criticized things. I was expecting to be given

> a project and do as instruct but, instead, they said, “Here is your idea. Let us go with it and see what happens.” It was a much more interesting experience for sure. (CU, MnDRIVE, Zoom interview, August 8, 2022)

The small size of each research group helped with a better atmosphere, mentioned by many participants, in terms of mutual learning (i.e., between mentors and mentees and among peers), receiving encouragement and suggestions, and spending more time together. Compared with traditional large classroom settings, the Victual Ep. program’s two research groups contained seven students each, and the size of the research teams of MnDRIVE projects were varied but smaller. Increased focus on individuals, individualized instructions, concentrated resources to support personal interests, close relationships due to mutual understanding, and community-based workspace—all these features gave participants a distinctive learning experience that most had not experienced elsewhere. The sense of community (i.e., in the research group/team as well as in LS-NSSA) led to the change in participants’ science identity and the sense of belonging:

> I did see many different kinds of people, definitely not strictly one type of gender, race, ethnicity, or background, and I do think everyone had a very good opportunity and they had their own time [to do the research] … Before I was like, am I the right person for STEM? I was not really sure, but now I think I definitely do enjoy the science, technology, engineering, math area more than other fields. I do definitely feel more like I belong it. (EC, Virtual Ep., Zoom interview, July 13, 2022)

> I felt like I was there [in the academic program] but not really part of anything, I was just continuing what I was hoping to do. But now that I become a *Louis Stokes scholar*. I feel that I am part of something, and I can see other people that are also trying to do their own things like me and be able to relate to the troubles and challenges that they are going through as well. (EI, MnDRIVE, Zoom interview August 2, 2022)

LS-NSSA created a counterspace for URM STEM students where they felt comfortable and empowered to contribute to research and share creative ideas. It provided opportunities and resources for URM students with specific attention to the ones who had the least access to them from other venues. However, outside of this counterspace, students still faced oppression and inequities on campus and in society. As KT (August 3, 2022) shared:

> A lot of times, you do not feel certain careers or certain environments are for you, even being in this STEM environment, even having Black and Brown people around, people who are from different countries, I still felt that maybe I do not belong there. So, I think that just shows the strength of *social conditioning* and things that we grew up with them—things that we *live with in society*… The program experience is really helpful because that can *change your mindset, change what is you are accustomed to, and show you different ways, and then fund you to do it*… In my prospects, especially as a Black woman, I just feel that the world is against me to succeed. So, just having these small things could possibly help lead me to success… *To take advantage*—that is really a big thing.

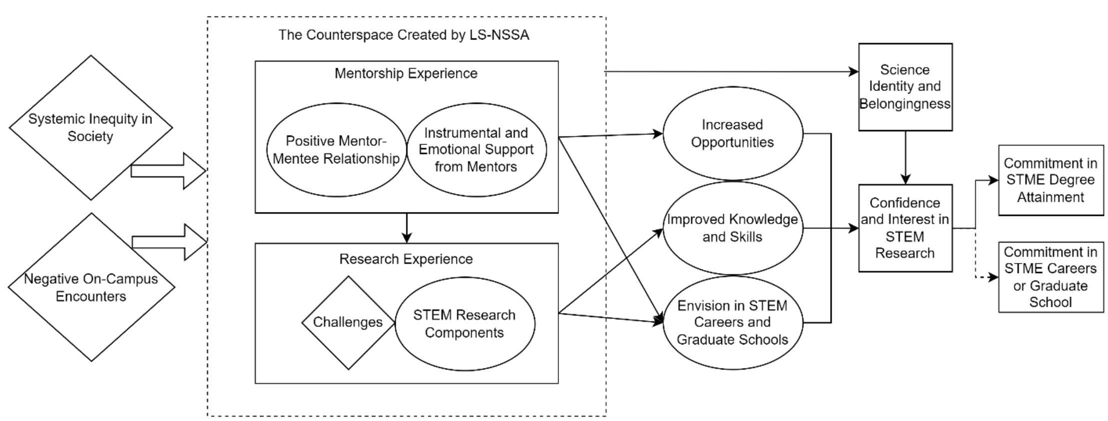
I was pretty much trying to take advantage of what I can have. (KT, MnDRIVE, Zoom interview, August 3, 2022)

Conclusion

Within a diverse and inclusive counterspace, participants perceived benefits from LS-NSSA's research mentorship programs (i.e., intellectually and emotionally) and were positively influenced by the research mentorship experiences. With increased opportunities in STEM fields, improved knowledge and skills in STEM research, and interactions with faculty and graduate students, there were positive changes in participants' confidence (i.e., self-efficacy), STEM interest, science identity development, and belongingness. These changes further promoted their commitment to STEM degree attainment. Some participants firmly committed more to postgraduate education in STEM disciplines and career pathways. Figure 12 displays the program theory of the research mentorship programs by integrating the SCCT framework and the contextual elements addressed through a CRT lens. Figure 13 shows the general theory of change elements addressed by the qualitative analysis (as shaded), focusing on the social-cognitive trajectory.

Figure 12

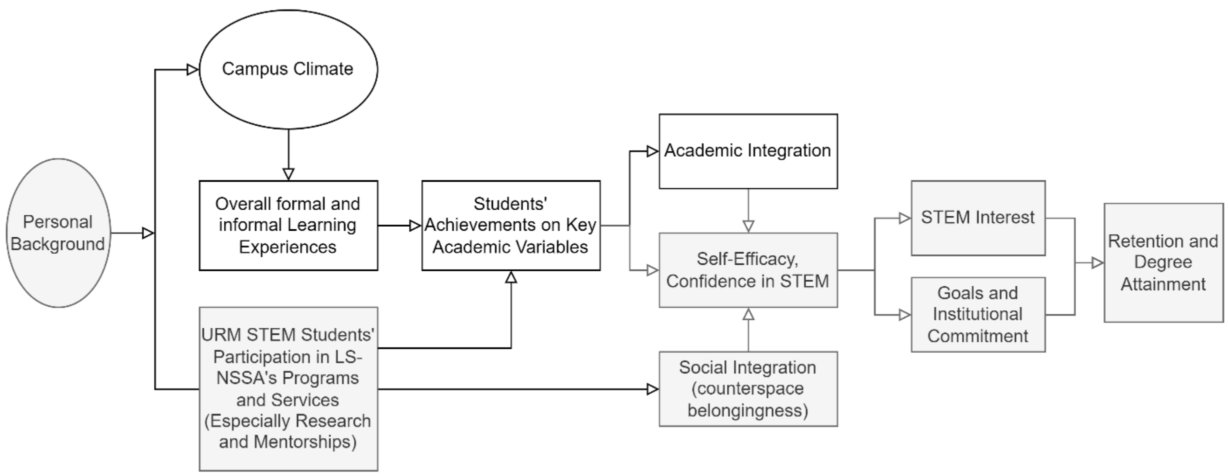
Sub-Theory of LS-NSSA's Research Mentorship Program



Note. (a) The solid arrows are relationships observed from sub-study 3; the dashed arrows (e.g., from "Confidence and Interest in STEM Research" to "Commitment in STEM Careers/Graduate School") have not been fully supported by the results. (b) The diamond-shaped items are negative factors in the trajectory; the rectangle items are first-level categories, indicating the main program elements and major outcomes; the oval items are second-level categories, indicating the intermediate elements contained in or generated from program activities.

Figure 13

Social Integration Trajectory Evaluated by Qualitative Analysis



Chapter Summary

The three research questions were answered by the three sub-studies, respectively. The results show that first-semester experience and academic outcomes influenced transfer students' graduation and retention in the three-year window and affected traditional students' graduation and retention in the four-year window. Through the comparisons between LS-NSSA participants and their counterparts (i.e., URM STEM non-participants and non-URM STEM students) on critical first-semester metrics and the first-year retention rate, the study shows that LS-NSSA made participants more academically prepared to pursue a bachelor's degree than URM students not affiliated with this program. Participants became competitive with White and Asian American STEM undergraduate students. Beyond the academic trajectory addressed by the quantitative analysis, the qualitative inquiry explains the social-cognitive trajectory through interviews with participants of two research mentorship programs. At LS-NSSA, URM STEM students had the opportunities and resources to enrich their undergraduate research experiences and strengthen their connections with STEM mentors. The increased confidence, interest, identity, and belongingness contributed to their commitment to STEM.

CHAPTER V. DISCUSSION AND IMPLICATIONS

This chapter interprets and discusses the results and findings. The current study investigated the effects of LS-NSSA participation on the retention and degree attainment of URM STEM students. A quantitative analysis investigated academic integration, and a qualitative analysis investigated social integration. The results are consistent with previous studies regarding factors and interventions that influence and support retention and degree attainment of URM STEM students. As part of LS-NSSA's internal evaluation, the results inform the effectiveness of this program by clarifying the key factors, comparing URM STEM students' academic outcomes with those of their counterparts, and describing the gains, challenges, and experiences of the target population from their culturally diverse perspectives. In this chapter, the main results and findings are reviewed, the contribution and limitations of the current study are discussed, and the implications for future research are proposed.

The Main Results and Findings

With reference to the hypotheses addressed in Chapter 1, the current study suggests that: (a) URM students' academic success in the STEM fields requires systemic effort, as it is associated with a range of distinctive factors (i.e., family background, individual academic background, individual social-cognitive factors, academic performance, campus climate, intervention and resources); (b) URM STEM students encounter systemic barriers and difficulties along their academic pathways, which are compounded by factors associated with their cultural identities; (c) undergraduate research and faculty mentorship have played critical roles in URM STEM students' learning experience and academic choices; (d) program mechanisms and their contextual factors can be understood through a mixed-methods theory-based evaluation approach; and (e) participants' reflections on their experience of the program is as essential a source of knowledge as statistical analysis for exploring, explaining, and evaluating the program from culturally diverse perspectives, primarily in relation to a critical race theory lens.

Significant Demographic and Academic Factors

Consistent with previous research, the results of the statistical analysis reported here suggest that URM STEM students' challenges and outcomes relate to (a) their demographic characteristics, such as their gender and family education background (Soria, 2014; Crisp & Nuñez, 2014; Dika & D'Amico, 2016; Wang, 2016; Belser et al., 2017; Wang et al., 2017), and (b) their academic performance, as measured by GPA (Rogulkin, 2011; Dika & D'Amico, 2016; Pedraza & Chen, 2022), especially during the first semester. The current study differentiates between transfer students and traditional students regarding factors affecting their retention and degree attainment across the three-year and four-year time frames. The results specifically address the importance of first-semester academic performance for transfer and traditional URM STEM students. Table 20 summarizes the direction of influential factors on transfer and traditional URM STEM students' retention and degree attainment, as generated from four models.

Table 20

Positive and Negative Factors Related to Transfer and Traditional URM STEM Students' Degree Attainment and Retention

Significant Factor	Transfer Student		Traditional Student	
	Degree Attainment	Retention	Degree Attainment	Retention
Demographic Factor				
First-Generation Status	(-)			

Gender (male)				(+)
On-Campus Housing Status at Entry			(+)	
Distance of Home to the Metro Area		(-)	(-)	(-)
Academic Factor				
Number of Transfer Credits	(+)	(+)	(+)	
First-Semester GPA		(+)	(+)	(+)
First-Semester Credit-Unit Pass Rate	(+)			(+)
Stop-Out Records	(-)		(-)	

Note. (a) The green and red colors are used to indicate the positive and negative factors, respectively. (b) The results are generated from four different models. The tables are used to visualize the results, not to compare results across models.

Demographic Factors. For transfer students, being the first college student in a household brings additional challenges and reduces their likelihood of graduating within the three-year window. This attribute deserves attention because a significant proportion of first-generation and transfer students have been engaged in LS-NSSA’s programs. A gender disparity has been observed among traditional students regarding their likelihood of retention, in that females experience more barriers than males. Traditional students are also influenced by their living conditions upon arrival. Relative to those who live off campus, on-campus residential options help enhance traditional students’ early engagement and benefit their academic outcomes, which has also been suggested by previous studies (LaNasa et al., 2007; Schudde, 2011; Graham et al., 2018).

The location of the permanent residence is associated with traditional students’ retention and degree attainment in the four-year window. Understandably, families play a critical role in traditional students’ personal development. Family members who are closer to campus can deliver needed financial, emotional, and physical support to first-year college traditional students, which promotes their general well-being and sense of belonging. For transfer students, who have a higher average age at entry, their own households may have more of an influence than that of their parents. Thus, the location of the permanent residence may not be related to transfer students' degree attainment in the three-year window. However, transfer students with residences outside of the state are less likely to be retained (if not graduated in time) relative to those who have permanent home addresses in the metropolitan area of the state. This indicates that family support is still important for those transfer students who experience more challenges on the way to graduation.

Academic Factors. The number of transfer credits is positively associated with transfer students' retention and degree attainment in the three-year window. Having more credits transferred from the previous institution appears to strengthen the knowledge base of transfer students in STEM studies and reduces their university course load. This assists transfer students to complete the degree requirement in the expected time frame. The number of transfer credits is associated with traditional students' degree attainment in the four-year window. Although the university allows students who enroll directly from high school to declare a major no later than the spring semester of the second year, an early determination to elect a STEM major and taking more related advanced placement (AP) courses in high school can accelerate traditional students' pace toward degree attainment. However, having taken fewer AP courses in high schools does not affect traditional students' retention in STEM education.

First-semester GPA is positively associated with traditional students' retention and degree attainment in four years. For traditional students who cannot graduate within four years of enrollment, a higher percentage of credit units passed in the first semester is associated with a greater likelihood of their retention in STEM. For transfer students, first-semester GPA is associated with retention among those who do not graduate in three years. A higher percentage of credit units passed in the first semester enhances the possibility of transfer students' degree attainment in three years.

The Importance of First-Semester Academic Performance. Regardless of how the factors described above are associated with transfer and traditional students' academic outcomes, statistically significant variables tend to describe first-semester performance, including first-semester GPA, first-semester credit-unit pass rate, on-campus residential status, and total number of transfer credits accepted by the STEM program at entry.

Previous studies identified transfer shock among transfer students in STEM disciplines when they encounter a different campus climate, altered achievement expectations, and a higher level of competitiveness (Reyes, 2011; Lakin & Elliott, 2016; Allen et al., 2022). Therefore, the first semester after the transfer is critical, and suggests a need for interventions to smooth transfer students' integration into their institution and support them academically, socially, and emotionally. The current study also shows that, for students of underrepresented racial/ethnic groups, first-semester academic performance is important for traditional students' choices and outcomes throughout their academic careers. Even though new high school students are not expected to claim their major directions of study at entry, their academic performances on first-semester courses tend, to a certain extent, to be connected with their paths and outcomes of the undergraduate learning experience in STEM disciplines.

LS-NSSA Enhances URM STEM Students' Academic Integration

LS-NSSA's influence on URM STEM students' academic integration is observed through group comparisons. The comparison between LS-NSSA's participants and their non-participant URM STEM counterparts seeks to control for differences attributable to demographic disparities. The comparison between LS-NSSA's participants and their non-URM (i.e., White and Asian American) STEM counterparts highlights these demographic disparities and seeks to examine whether LS-NSSA helps narrow the achievement gap. Compared with White and Asian-American STEM students, although LS-NSSA's participants have a lower first-semester GPA (i.e., 3.14 vs. 3.31), this difference is *not* statistically significant. In addition, LS-NSSA's participants' first-year retention rate is one percent higher than that of non-URM students (i.e., 96% vs. 95%) and significantly higher than URM STEM non-participants (i.e., 96% vs. 79%). These results are an essential recognition of LS-NSSA's effort because the first-year retention is a critical step toward the consistent registration and degree attainment for URM STEM students. It appears that the achievement gap between URM STEM students and historically privileged groups may be narrowed through effective programs that improve participants' first-semester academic performance.

LS-NSSA provides early intervention to support target populations' learning experiences. This involves financial support (e.g., the First Year Scholar Program that provides a maximum $1,000 stipend for participation, as well as conference travel support), research opportunities (e.g., faculty-sponsored research projects and research mentorship programs), and professional development workshops. These programs have engaged first-year traditional URM STEM students and newly enrolled transfer students, offering resources and opportunities that those students may not be able to obtain from other sources.

The qualitative results suggest the benefits that URM STEM students can obtain from their early affiliation with LS-NSSA. Because the research mentorship programs are open to students with entry-level research knowledge and skills, first-year students and students who intend to transfer to the university are all welcomed and engaged, either virtually or in a laboratory setting. This allows participants to understand STEM research and career pathways, improve their STEM knowledge and skills, and build connections with faculty mentors. These changes are positively related to their academic performance, intellectual development, and learning experiences in academic settings.

LS-NSSA Enhances URM STEM Students' Social Integration

Tinto (1975, 1997) defined *social integration* as students' positive relationship with peers, affiliation with on-campus organizations, and involvement in extracurricular activities. The current study uses this concept, adding students' perceived diversity, equity, and inclusion, which are presumed to influence historically underrepresented populations' engagement, belongingness, and success in the higher education system.

As previous studies have suggested, social-cognitive factors such as self-efficacy/confidence, STEM interest, goal commitment, and belongingness are critical elements in a program's mechanisms targeting URM STEM students' retention and degree attainment (e.g., Lent et al., 1994; Tinto, 1997; Estrada et al., 2011; Potvin & Hazari, 2013; Byars-Winston et al., 2015; Syed et al., 2019; Wang et al., 2017; Allen et al., 2022). The current study provides additional evidence to support these findings. At LS-NSSA, positive mentorship experiences and undergraduate research experiences appear to reinforce each other in a culturally sensitive environment. The positive mentor-mentee interactions and research practices led to increased research opportunities, expanded academic and social networks, a sense of belonging, science identity, confidence, and learning interests. They contributed to a stronger on-campus affiliation across URM STEM participants, thus generating a more substantial commitment to bachelor's degree attainment in STEM disciplines (see Figure 12 in Chapter IV).

LS-NSSA provided mentor and mentee training at the beginning of the research mentorship programs to enhance positive mentor-mentee interactions. The topics covered culturally appropriate instruction, learning, communication, and problem-solving strategies. The training helped address the importance of mutual learning in diverse contexts. Mentors were aware of potential mentor-mentee disparities related to cultural identities and cultural values. Although program participants did not expect that mentors would share their demographic backgrounds, positive experiences were usually associated with open, supportive, flexible, responsible, and approachable mentorships. A key mentor quality reported by students was caring and balancing of mentor expectations with their mentees' needs. Students in advanced courses may lack adequate experience and skills to conduct scientific research. This is especially common among URM students studying in STEM disciplines, taking into account the difficulties and challenges that they have experienced in higher education. They need programs that allow them to be comfortable despite gaps in knowledge, encouraging them to build up their resumes, and giving them a chance to take action.

Participants' social integration is promoted by positive relationships with mentors, but more importantly, perceived levels of acceptance, respect, comfort, and representation. The participants described having experienced microaggressions, disproportionate representation, and limited exposure or accessibility to resources on campus, which exacerbated barriers to their advancement in STEM. However, LS-NSSA creates a counterspace for URM STEM students where they can experience diversity and inclusion, stronger peer connection, culturally diverse social networks, empowerment, and autonomy in learning. Being systemically underserved, URM STEM students require greater attention and resources to promote academic success. LS-NSSA plays this role, complementing the unbalanced financial, research, and mentoring resources for URM STEM students. This counterspace, to a certain extent, balances negative encounters with the larger system and promotes positive URM STEM student outcomes.

The Contribution of the Current Evaluative Study

The current study is a part of the internal evaluation of LS-NSSA. It contributed an improved understanding of how LS-NSSA works and investigated the program's theory of change and the influence of LS-NSSA's activities. It also contributed to research on evaluation by applying a theory-based evaluation approach to a culturally sensitive intervention and illustrated how qualitative and quantitative evidence can be integrated into an explanation of a program's theory of change.

The Value and Use of the Evaluation Results

Program evaluation is applied research. Evaluators systematically collect information on an evaluand's inputs, activities, characteristics, outputs, and outcomes to determine its merit and worth (Scriven, 1991), support decision-making (Weiss, 1988), and promote organizational learning and improvement (Patton, 2008). A theory-based evaluation approach serves these purposes by understanding the impact of the evaluand and specifying its route to the desired goals (Chen, 1990; Coryn, et al., 2011). The current theory-based evaluative study generated an understanding of LS-NSSA and promoted the utilization of evaluation results.

Knowledge Generation. After 15 years of program activities, over 700 URM LS-NSSA undergraduates obtain STEM bachelor's degrees every year across the Alliance. The current study represents the first theory-based internal evaluation investigating linkages between program inputs/activities and participants' retention and graduation. The current study utilized Tinto's (1975, 1993) student dropout model, in combination with Lent et al.'s (1994) Social Cognitive Career Theory framework, to investigate student academic and social integration trajectories. This integrated model was adapted to the context of LS-NSSA's program, accounting for DEI factors (i.e., campus climate and counterspace), and probing the program's influence on student trajectories.

The current study contributed an investigation of LS-NSSA's theory of change and sub-theory of research mentorship programs. Rather than focusing on inputs and outputs via a logic model, the program theory was investigated with respect to changes in URM STEM students' academic performance and psychology as a consequence of program participation. This theory of change focused on key predictors and moderating factors and their relationships with the key outcomes of URM STEM students' retention and degree attainment. Attention focused on investigation of LS-NSSA research mentorship programs to explain the influence of specific elements (i.e., mentorship and research experience) on one of the key moderating factors—participants' social integration. The analysis illustrated how program features may be explored and explained through a theory-based evaluation approach. The results provide guidance regarding program elements that contribute to improved outcomes.

Evaluation Utilization. The utilization of evaluation results refers to how "an evaluation and information from the evaluation impacts the program that is being evaluated" (Alkin & Taut, 2003, p. 1). LS-NSSA's internal evaluation plan articulated in the program's 2016 grant proposal specified monitoring activities and objective-focused sub-studies (see Appendix B). These internal efforts sought to provide continuous support for program implementation and improvement. The current study sought to explain the impact of the LS-NSSA program and to inform the next five years of implementation and evaluation.

The results clarified the importance of URM STEM students' first-semester learning outcomes. The results reinforced LS-NSSA's focus on first-year scholars. Additional efforts to promote first-semester academic performance through study groups, academic tutoring, and peer mentoring may be beneficial. The results identified qualities of mentorships that promote URM STEM students' success.

The results suggest that mentor training should focus on providing developmentally appropriate emotional support, maintaining consistent and frequent interactions, being flexible and open, and promoting URM student empowerment.

The results clarified the paths by which LS-NSSA improves URM STEM students' social integration and academic integration. LS-NSSA may utilize this information to institutionalize program activities that promote social and academic integration.

Finally, the results may be useful to LS-NSSA as a model of theory-based evaluation that may be used as an evaluation approach to investigate additional LS-NSSA program elements and activities.

The Application of Theory-Based Evaluation in Internal Evaluation

The evaluation that was conducted was an internal evaluation. Typically, an internal evaluation focuses on planning, monitoring, and judging the process of program implementation and service delivery. Internal evaluators typically focus on utilization studies, satisfaction assessment, quality assurance activities, cost-benefit analysis, and self-study practices (Love, 1991).

In contrast, the current study contributed a theory-based internal evaluation. A theory-based evaluation approach is well-aligned with LS-NSSA's needs. First, LS-NSSA is rationalized in terms of theory that is directly applicable to URM STEM students (i.e., Tinto's model, 1993) but the theory had not previously been tested via a program evaluation at LS-NSSA. Second, LS-NSSA has been recognized as a mature program aiming for broad statewide influence across a network of higher education institutions. The current study sought to articulate and investigate the theory of change to inform program improvement.

A theory-based internal evaluation approach seeks to balance scientific credibility with practical recommendations to support decision-making and program improvement. The current study illustrated how a theory-based evaluation approach can be utilized to investigate relationships among program elements, moderating factors, and outcomes, and to identify problematic linkages that require attention.

The Integration of Mixed Methods in the Program Context. In a theory-based evaluation, the evaluand's theory of change is the theoretical framework of the evaluation, and the mixed-methods design is embedded within it. The current study illustrated how quantitative and qualitative evidence generated from an integrated mixed-methods approach can be used to develop a theory of change. The results informed an understanding of URM STEM students' challenges, focusing on factors, attributes, characteristics, and experiences that influenced persistence, retention, and progress toward graduation. DEI issues were addressed via the theoretical framework and contextualized interpretation of the results. The results suggest the importance of LS-NSSA's influence on the campus climate experienced by URM STEM students.

Limitations

The current study has three main limitations. First, its scope was limited to a single institutional source of data regarding the Alliance. Although the University of Minnesota has the largest portion of program participants and two-thirds of URM STEM graduates, the representativeness of this study would be improved by engaging Alliance-wide samples. The University of Minnesota is a predominantly White institution (PWI) with 65% White students. Its campus climate may be different from some Alliance partners serving more diverse populations. Partners with lower percentages of White student enrollment include Augsburg University and Carleton College (42% and 58%, respectively; College Factual, n.d.). Partners such as the Bemidji State University have been promoting Indigenous STEM students' engagement and outcomes, which was disproportionately represented in the current samples. The existing Alliance-wide evaluation system tracks institutional outcomes, such as URM STEM students' enrollment and degree attainment information. Beyond that, program participants' academic information, engagement information, and experiential knowledge have yet to be shared across the Alliance. These factors prevented an in-depth Alliance-wide investigation.

Second, the current theory of change addresses the trajectories of URM STEM students enrolled at four-year institutions, including transfer students and traditional students. However, the Alliance includes several community colleges. These are feeder institutions that connect transfer students with four-year partners. Students pursuing transfer opportunities and STEM-related careers in community colleges are also the targets of LS-NSSA. Increased enrollment and retention of transfer students would contribute to overall outcomes in the URM population in STEM. The program theory does not provide information regarding program elements that influence these students' choices and actions.

Third, in terms of evaluation methods, the theory-based evaluation approach addressed the links between the program elements and the outcomes, but these links were associative, not causal. The interpretation of the theory of change was contextualized in the current program, depending on program implementation and program-system interactions. It may not work for similar programs in other contexts. Applying the evaluation results to other situations would require additional investigation. In addition, the current mixed methods approach was partially triangulated within the theoretical framework. For instance, social integration involves various social-cognitive factors that should be tested through statistical modeling to improve confidence in the results. Student academic trajectories are likely influenced by environmental factors related to campus climate and informal learning experiences, which have yet to be probed.

Future Studies

With respect to these limitations of this study and the gap of knowledge that remains, future evaluative studies may wish to focus on three aspects.

First, it is necessary to enlarge the scope of the evaluative study to engage Alliance-wide data analysis in relation to the key elements in the theory of change. The dataset may also involve the pre-entry attributes of the target population (e.g., course patterns in high schools or community colleges, pre-entry cumulative GPAs) that may be associated with students' first-semester experience. This information can also support an improved understanding of barriers in transfer pathways.

Second, the internal evaluator may undertake a quantitative analysis using self-reported survey data to validate the influence of the program on URM STEM students' social integration trajectory. The author and colleagues developed a participation survey and disseminated it across the Alliance over the past two years. The response rate was low during the pandemic but has increased with Alliance-wide efforts. The survey contains constructs that measure social-cognitive factors and students' experiences with distinctive program features. Data collection is ongoing. The sample is expected to cover participants from partnering institutions, enhancing Alliance-wide evidence regarding the theory of change.

Third, additional evaluation activities may probe the program's influence on campus climate, with a focus on cultural aspects. This study has touched on this topic by discussing students' perspectives and experiences on campus and in the program. However, it is difficult to examine the assumptions and values embedded in the culturally diverse community where LS-NSSA operates. Campus climate is treated as an element of the context, rather than a key element of the program's mechanism. Further studies are needed to investigate the implications with respect to program institutionalization.

References

Adedokun, O. A., Parker, L. C., Childress, A., Burgess, W., Adams, R., Agnew, C. R., Leary, J., Knapp, D., Shields, C., Lelievre, S. & Teegarden, D. (2014). Effect of time on perceived gains from an undergraduate research program. *CBE—Life Sciences Education*, *13*(1), 139-148. http://dx.doi.org/10.1187/cbe.13-03-0045

Allen, D., Dancy, M., Stearns, E., Mickelson, R., & Bottia, M. (2022). Racism, sexism and disconnection: contrasting experiences of Black women in STEM before and after transfer from community college. *International Journal of STEM Education*, *9*(1), 1-21. http://dx.doi.org/10.1186/s40594-022-00334-2

Anderson, E., & Kim, D. (2006). *Increasing the graduation of minority students in science and technology [4th in a Series: The Unfinished Agenda: Ensuring graduation for students of color]*. Washington, DC: American Council on Education.

Archer, M., Bhaskar, R., Collier, A., Lawson, T., & Norrie, A. (2013). *Critical realism: Essential readings.* Routledge.

Astbury, B., Frans, L., & Leeuw, F. L. (2010). Unpacking black boxes: Mechanisms and theory building in evaluation, *American Journal of Evaluation, 31* (3), 363-381. http://dx.doi.org/10.1177/1098214010371972

Alkin, M. C., & Taut, S. M. (2003). Unbundling evaluation use. *Studies in Educational Evaluation*, *29*, 1-12. http://dx.doi.org/10.1016/S0191-491X(03)90001-0

Atkins, K., Dougan, B. M., Dromgold-Sermen, M. S., Potter, H., Sathy, V., & Panter, A.

T. (2020). “Looking at Myself in the Future”: how mentoring shapes scientific identity for STEM students from underrepresented groups. *International Journal of STEM Education*, *7*(1), 1-15. http://dx.doi.org/10.1186/s40594-020-00242-3

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191. http://dx.doi.org/10.1037/0033-295X.84.2.191

Bandura, A., & National Inst of Mental Health. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall, Inc.

Bhaskar, R. (2008). *A Realist Theory of Science.* New York: Routledge.

Bell, L. A. (2003). Telling tales: What stories can teach us about racism. *Race, Ethnicity and Education*, *6*(1), 3-28. http://dx.doi.org/10.1080/1361332032000044567

Bell, L. A. (2009). Learning through story types about race and racism. *Social justice education: Inviting faculty to transform their institutions*, 26-41.

Belser, C. T., Prescod, D. J., Daire, A. P., Dagley, M. A., & Young, C. Y. (2017). Predicting Undergraduate Student Retention in STEM Majors Based on Career Development Factors. *The Career Development Quarterly*, *65*(1), 88-93. https://doi.org/10.1002/cdq.12082

Betts, J. (2013). Aid effectiveness and governance reforms: Applying realist principles to a complex synthesis across varied cases. *Evaluation, 19*(3), 249-268. http://dx.doi.org/10.1177/1356389013493840

Bhaskar, R., Collier, A., Lawson, T., & Norrie, A. (1998). Critical realism. In *Proceedings of the Standing Conference on Realism and Human Sciences, Bristol, UK (Vol. 4).*

Birckmayer, J. D., & Weiss, C. H. (2000). Theory-based evaluation in practice: what do we learn?. *Evaluation review*, *24*(4), 407-431. http://dx.doi.org/10.1177/0193841X0002400404

Blake, R. A., Liou-Mark, J., & Chukuigwe, C. (2013). An effective model for enhancing underrepresented minority participation and success in geoscience undergraduate research. *Journal of Geoscience Education*, *61*(4), 405-414. http://dx.doi.org/10.5408/12-417.1

Blake-Beard, S., Bayne, M. L., Crosby, F. J., & Muller, C. B. (2011). Matching by race and gender in mentoring relationships: Keeping our eyes on the prize. *Journal of Social issues*, *67*(3), 622-643. http://dx.doi.org/10.1111/j.1540-4560.2011.01717.x

Brousselle, A., & Buregeya, J.-M. (2018). Theory-based evaluations: Framing the existence of a new theory in evaluation and the rise of the 5th generation. *Evaluation*, *24*(2), 153–168. http://dx.doi.org/10.1177/1356389018765487

Brown, N. N., Velarde, J., & Mascaro, D. J. (2016). Using Peer Mentoring to Enhance Student Experience and Increase Retention in Mechanical Engineering. In *ASEE Annual Conference & Exposition*. http://dx.doi.org/10.18260/p.27157

Brownell, S. E., Hekmat-Scafe, D. S., Singla, V., Chandler Seawell, P., Conklin Imam, J. F., Eddy, S. L., Stearns, T., Cyert, M. S., & Hewlett, J. (2015). A high-enrollment course-based undergraduate research experience improves student conceptions of scientific thinking and ability to interpret data. *CBE—Life Sciences Education*, *14*(2), ar21.

https://doi.org/10.1187/cbe.14-05-0092

Byars-Winston, A., Rogers, J., Branchaw, J., Pribbenow, C., Hanke, R., & Pfund, C. (2016). New measures assessing predictors of academic persistence for historically underrepresented racial/ethnic undergraduates in science. *CBE-Life Sciences Education*, *15*(3), ar32. http://dx.doi.org/10.1187/cbe.16-01-0030

Cain, E. (2017). *Understanding the experiences of under-represented minority transfer students in STEM disciplines at the University of Minnesota Twin Cities*. Unpublished manuscript.

Campbell, D. T. and Stanley, J. C. (1966) *Experimental and quasi-experimental designs for research. Skokie*, IL: Rand McNally.

Carpi, A., Ronan, D. M., Falconer, H. M., Boyd, H. H., & Lents, N. H. (2013). Development and Implementation of Targeted STEM Retention Strategies at a Hispanic-Serving Institution. *Journal of Hispanic Higher Education*, *12*(3), 280–299. http://dx.doi.org/10.1177/1538192713486279

Carpi, A., Ronan, D. M., Falconer, H. M., & Lents, N. H. (2017). Cultivating minority scientists: Undergraduate research increases self-efficacy and career ambitions for underrepresented students in STEM. *Journal of Research in Science Teaching*, *54*(2), 169-194. http://dx.doi.org/10.1002/tea.21341

Carroll, M. A., & Barnes, E. F. (2015). Strategies for enhancing diverse mentoring relationships in STEM fields. *International Journal of Evidence Based Coaching and Mentoring*, *13*(1), 58-69.

Case, A., & Hunter, C. (2012). Counterspaces: A Unit of Analysis for Understanding the

Role of Settings in Marginalized Individuals' Adaptive Responses to Oppression. *American Journal of Community Psychology*, *50*(1-2), 257-270. http://dx.doi.org/10.1007/s10464-012-9497-7

Castleman, B. L., Long, B. T., & Mabel, Z. (2018). Can financial aid help to address the growing need for STEM education? The effects of need-based grants on the completion of science, technology, engineering, and math courses and degrees. *Journal of Policy Analysis and Management*, *37*(1), 136-166. http://dx.doi.org/10.1002/pam.22039

Castro, E. L. (2014). "Underprepared" and "at-risk": Disrupting deficit discourses in undergraduate STEM recruitment and retention programming. *Journal of Student Affairs Research and Practice*, *51*(4), 407-419. http://dx.doi.org/10.1515/jsarp-2014-0041

Chang, M. J., Sharkness, J., Hurtado, S., & Newman, C. B. (2014). What matters in college for retaining aspiring scientists and engineers from underrepresented racial groups. *Journal of Research in Science Teaching*, *51*(5), 555–580. http://dx.doi.org/10.1002/tea.21146

Charleston, L. J., George, P. L., Jackson, J. F., Berhanu, J., & Amechi, M. H. (2014). Navigating underrepresented STEM spaces: Experiences of Black women in US computing science higher education programs who actualize success. *Journal of Diversity in Higher Education*, *7*(3), 166. http://dx.doi.org/10.1037/a0036632

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis.* Sage.

Charmaz, K. (2014). *Constructing grounded theory*. Sage.

Chemers, M. M., Zurbriggen, E. L., Syed, M., Goza, B. K., & Bearman, S. (2011). The role of efficacy and identity in science career commitment among underrepresented minority students. *Journal of Social Issues, 67*(3), 469-491. http://dx.doi.org/10.1111/j.1540-4560.2011.01710.x

Chen, H. T. (1989). The conceptual framework of the theory-driven perspective. *Evaluation and Program Planning, 12*, 391-396. http://dx.doi.org/10.1016/0149-7189(89)90057-8

Chen, H. T. (1990). *Theory-driven evaluations*. Sage.

Chen, H. T. (2006). A theory-driven evaluation perspective on mixed methods research. *Research in the Schools, 13*(1), 75-83.

Chen, H. T. (2012). Theory-driven evaluation: Conceptual framework, application and advancement. In *Evaluation of programs and projects for a democratic culture,* 17-40. Springer. http://dx.doi.org/10.1007/978-3-531-19009-9_2

Chen, H. T. (2014). *Practical program evaluation: Theory-driven evaluation and the integrated evaluation perspective.* Sage.

Chen, H. T., & Rossi, P. H. (1980). The multi-goal, theory-driven approach to evaluation: A model linking basic and applied social science. *Social forces, 59*(1), 106-122. http://dx.doi.org/10.2307/2577835

Chen, H. T., & Rossi, P. H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation review, 7*(3), 283-302. http://dx.doi.org/10.1177/0193841X8300700301

Chen, H. T., & Rossi, P. H. (1987). The theory-driven approach to validity. *Evaluation* *and program planning, 10*(1), 95-103. http://dx.doi.org/10.1016/0149-7189(87)90025-5

Chen, H. T., & Rossi, P. H. (1989). Issues in the theory-driven perspective. *Evaluation and program planning*, *12*(4), 299-306. http://dx.doi.org/10.1016/0149-7189(89)90046-3

Chen, X. (2013). STEM Attrition: College Students' Paths into and out of STEM Fields. Statistical Analysis Report. NCES 2014-001. *National Center for Education Statistics*.

College Factual. (n.d.). *Find your college at CollegeFactual.com: Demographics & Diversity Report*. https://www.collegefactual.com/

Cooper, K. M., Gin, L. E., Akeeh, B., Clark, C. E., Hunter, J. S., Roderick, T. B., et al. (2019). Factors that predict life sciences student persistence in undergraduate research experiences. *PLoS ONE, 14*(8). http://dx.doi.org/10.1371/journal.pone.0220186

Coryn, C. L., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009. *American journal of Evaluation, 32*(2), 199-226. http://dx.doi.org/10.1177/1098214010389321

Cotts, E., Fang, J., Jones, W., Klotzkin, D., Myers, G., & White, B. (2016, October). A STEM program focused on transfer student success at Binghamton University. In *Frontiers in Education Conference (FIE), 2016 IEEE* (1-4). IEEE. http://dx.doi.org/10.1109/FIE.2016.7757551

Creswell, J. W. (1999). Mixed-method research: Introduction and application. In *Handbook of educational policy* (455-472). Academic press.

Creswell, J. W. (2010). Mapping the developing landscape of mixed methods research. *SAGE handbook of mixed methods in social & behavioral research*, *2*, 45-68. http://dx.doi.org/10.4135/9781506335193.n2

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Creswell J. W., Plano Clark V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage.

Crisp, G., & Nuñez, A. M. (2014). Understanding the racial transfer gap: Modeling underrepresented minority and nonminority students' pathways from two-to four-year institutions. *The Review of Higher Education*, *37*(3), 291-320. http://dx.doi.org/10.1353/rhe.2014.0017

Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., Walker, D. F., & Weiner, S. S. (1980). *Toward reform of program evaluation (p. 3).* San Francisco: Jossey-Bass.

Dahler-Larsen, P. (2011). *The Evaluation Society*. Stanford University Press.

DeAngelo, L., Franke, R., Hurtado, S., Pryor, J. H., & Tran, S. (2011). *Completing college: Assessing graduation rates at four-year institutions*. UCLA Higher Education Research Institute.

Denzin, N. K. (2012). Triangulation 2.0. *Journal of mixed methods research, 6*(2), 80-88. http://dx.doi.org/10.1177/1558689812437186

Damkaci, F., Braun, T. F., & Gublo, K. (2017). Peer mentor program for the general chemistry laboratory designed to improve undergraduate STEM retention. *Journal of Chemical Education, 94*(12), 1873-1880. http://dx.doi.org/10.1021/acs.jchemed.7b00340

Davidson, E. J. (2004). *Evaluation methodology basics: The nuts and bolts of sound evaluation.* Sage.

Delgado, R., & Stefancic, J. (1993). Critical race theory: An annotated bibliography. *Virginia Law Review*, 461-516. http://dx.doi.org/10.2307/1073418

Delgado, R., & Stefancic, J. (2017). *Critical race theory: An introduction* (Vol. 20). NYU Press.

Dika, S. L., & D'Amico, M. M. (2016). Early experiences and integration in the persistence of first-generation college students in STEM and non-STEM majors. *Journal of Research in Science Teaching*, *53*(3), 368-383. http://dx.doi.org/10.1002/tea.21301

Dixson, A. D., & Rousseau, C. K. (2014). And we are still not saved: Critical race theory in education ten years later. In *Critical race theory in education* (45-68). Routledge. http://dx.doi.org/10.4324/9781315870557-8

Drew, J. C., Galindo-Gonzalez, S., Ardissone, A. N., & Triplett, E. W. (2016). Broadening participation of women and underrepresented minorities in STEM through a hybrid online transfer program. *CBE—Life Sciences Education*, *15*(3), ar50. http://dx.doi.org/10.1187/cbe.16-01-0065

D’Amico, M. M., Dika, S. L., Elling, T. W., Algozzine, B., & Ginn, D. J. (2014). Early integration and other outcomes for community college transfer students. *Research in Higher Education*, *55*(4), 370-399. http://dx.doi.org/10.1007/s11162-013-9316-5

Eby, L. T., McManus, S., Simon, S. A., & Russell, J. E. A. (2000). An examination of negative mentoring experiences from the protégé's perspective. *Journal of Vocational Behavior*, *57*(1), 42-61.

Eby L. T., Rhodes J. E., Allen T. D. (2007). Definition and evolution of mentoring. In Allen T. D., Eby L. T. (Eds.), *The Blackwell handbook of mentoring*, 7–20. http://dx.doi.org/10.1002/9780470691960.ch2

Eby, L. T. D. T., Allen, T. D., Hoffman, B. J., Baranik, L. E., Sauer, J. B., Baldwin, S., Morrison, M. A., Kinkade, K. M., Maher, C. P. & Evans, S. C. (2013). An interdisciplinary meta-analysis of the potential antecedents, correlates, and consequences of protégé perceptions of mentoring. *Psychological bulletin*, *139*(2), 441. https://doi.org/10.1037/a0029279

Eris, Ö., Chachra, D., Chen, H., Rosca, C., Ludlow, L., Sheppard, S., & Donaldson, K. (2007, June 24–27). *A preliminary analysis of correlates of engineering persistence: results from a longitudinal study*. Paper presented at the American Society for Engineering Education Annual Conference, Honolulu, Hawaii. http://dx.doi.org/10.18260/1-2--2781

Espinosa, L. L. (2011). Pipelines and pathways: Women of color in undergraduate STEM majors and the college experiences that contribute to persistence. *Harvard Education Review, 81*(2), 209–240. http://dx.doi.org/10.17763/haer.81.2.92315ww157656k3u

Estrada, M., Burnett, M., Campbell, A. G., Campbell, P. B., Denetclaw, W. F., Gutiérrez, C. G., Hurtado, S., John, G. H., Matsui, J., McGee, R., … & Zavala, M. (2016). Improving underrepresented minority student persistence in STEM. *CBE—Life Sciences Education*, *15*(3), es5. https://doi.org/10.1187/cbe.16-01-0038

Estrada, M., Hernandez, P. R., & Schultz, P. W. (2018). A longitudinal study of how quality mentorship and research experience integrate underrepresented minorities into STEM careers. *CBE—Life Sciences Education*, *17*(1), ar9. http://dx.doi.org/10.1187/cbe.17-04-0066

Estrada, M., Woodcock, A., Hernandez, P. R., & Schultz, P. W. (2011). Toward a model of social influence that explains minority student integration into the scientific community. *Journal of Educational Psychology*, *103*(1), 206. http://dx.doi.org/10.1037/a0020743

Feder, T. (2019). Ongoing mentorship works for retaining minorities in STEM. *Physics today*, *72*(10), 24-27. http://dx.doi.org/10.1063/PT.3.4315

Flores, L. Y., Navarro, R. L., Lee, H. S., Addae, D. A., Gonzalez, R., Luna, L. L., Jacquez, R., Cooper, S. & Mitchell, M. (2014). Academic satisfaction among Latino/a and White men and women engineering students. *Journal of Counseling Psychology*, *61*(1), 81. https://doi.org/10.1037/a0034577

Fouad, N. A., & Santana, M. C. (2017). SCCT and underrepresented populations in STEM fields: Moving the needle. *Journal of Career Assessment*, *25*(1), 24-39. http://dx.doi.org/10.1177/1069072716658324

Gayles, J. G., & Ampaw, F. (2014). The impact of college experiences on degree completion in STEM fields at four-year institutions: Does gender matter?. *The Journal of Higher Education*, *85*(4), 439-468. http://dx.doi.org/10.1353/jhe.2014.0022

Ghee, M., Keels, M., Collins, D., Neal-Spence, C., & Baker, E. (2016). Fine-tuning summer research programs to promote underrepresented students' persistence in the STEM pathway. *CBE—Life Sciences Education*, *15*(3), ar28. http://dx.doi.org/10.1187/cbe.16-01-0046

Graham, P. A., Hurtado, S. S., & Gonyea, R. M. (2018). The benefits of living on campus: Do residence halls provide distinctive environments of engagement?. *Journal of Student Affairs Research and Practice*, *55*(3), 255-269. http://dx.doi.org/10.1080/19496591.2018.1474752

Green, A. M., Brand, B. R., & Glasson, G. E. (2019). Applying actor–network theory to identify factors contributing to non-persistence of African American students in STEM majors. *Science Education*, *103*(2), 241-263.

Greene, J. C. (2006). Evaluation, democracy, and social change. *The Sage handbook of evaluation*, 118-140. http://dx.doi.org/10.4135/9781848608078.n5

Greenhalgh, T., Wong, G., Jagosh, J., Greenhalgh, J., Manzano, A., Westhorp, G., & Pawson, R. (2015). Protocol—the RAMESES II study: developing guidance and reporting standards for realist evaluation. *BMJ open*, *5*(8), e008567. http://dx.doi.org/10.1136/bmjopen-2015-008567

Guba, E. G., & Lincoln, Y. S. (1981). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches.* Jossey-Bass.

Guest, G., MacQueen, K. M., & Namey, E. E. (2011). *Applied thematic analysis*. Sage

Hanauer, D. I., Graham, M. J., & Hatfull, G. F. (2016). A measure of college student persistence in the sciences (PITS). *CBE—Life Sciences Education*, *15*(4), ar54. http://dx.doi.org/10.1187/cbe.15-09-0185

Harper, S. R. (2012). Race without racism: How higher education researchers minimize racist institutional norms. *Review of Higher Education*, *36*(1, Supplement), 9-29. http://dx.doi.org/10.1353/rhe.2012.0047

Hernández, E. (2016). Utilizing critical race theory to examine race/ethnicity, racism, and power in students' development theory and research. *Journal of College Student Development*, *57*(2), 168-180. http://dx.doi.org/10.1353/csd.2016.0020

Hernández, P. R., Estrada, M., Woodcock, A., & Schultz, P. W. (2017). Protégé perceptions of high mentorship quality depend on shared values more than on demographic match. *The Journal of Experimental Education*, *85*(3), 450-468. http://dx.doi.org/10.1080/00220973.2016.1246405

Higgins, T. B., Brown, K. L., Gillmore, J. G., Johnson, J. B., Peaslee, G. F., & Stanford, D. J. (2011). Successful student transitions from the community college to the four-year college facilitated by undergraduate research. *CUR Quarterly*, *31*(3), 16.

Higgins, M. C., & Kram, K. E. (2001). Reconceptualizing mentoring at work: A developmental network perspective. *Academy of management review*, *26*(2), 264-288. http://dx.doi.org/10.2307/259122

Higgins, M. C., & Thomas, D. A. (2001). Constellations and careers: Toward understanding the effects of multiple developmental relationships. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational* *and Organizational Psychology and Behavior*, *22*(3), 223-247. http://dx.doi.org/10.1002/job.66

Hiraldo, P. (2010). The role of critical race theory in higher education. *The Vermont Connection, 31*(1), 7.

Holma, K., & Kontinen, T. (2011). Realistic evaluation as an avenue to learning for development NGOs. *Evaluation*, *17*(2), 181-192. http://dx.doi.org/10.1177/1356389011401613

Hornickel, A. (2016). *NSSA 2017-2022 grant application narratives*. Unpublished internal document.

Howe, K. R. (2003). Closing methodological divides. *Toward democratic educational research* (Vol. 11). Springer Science & Business Media.

Hurtado, S., Newman, C. B., Tran, M. C., & Chang, M. J. (2010). Improving the rate of success for underrepresented racial minorities in STEM fields: Insights from a national project. *New Directions for Institutional Research, 2010*(148), 5-15. http://dx.doi.org/10.1002/ir.357

IBM Corp. (2017). *IBM SPSS Statistics for Windows (Version 25.0)* [Computer software]. IBM Corp.

Jackson, D. L. (2013). A balancing act: Impacting and initiating the success of African American female community college transfer students in STEM into the HBCU environment. *The Journal of Negro Education*, *82*(3), 255-271.

Jackson, D. L., Starobin, S. S., & Laanan, F. S. (2013). The shared experiences:

Facilitating successful transfer of women and underrepresented minorities in STEM fields. *New Directions for Higher Education*, *2013*(162), 69-76. http://dx.doi.org/10.1002/he.20058

Jagosh, J., Tilley, N., & Stern, E. (2016). Realist Evaluation at 25: Cumulating knowledge, advancing debates and innovating methods. *Evaluation*, *22*(3), 267-269. http://dx.doi.org/10.1177/1356389016656502

Jamil, M. G. (2018). Technology-enhanced teacher development in rural Bangladesh: A critical realist evaluation of the context. *Evaluation and program planning*, *69*, 1-9. http://dx.doi.org/10.1016/j.evalprogplan.2018.04.002

Jehangir, R. R., Stebleton, M. J., & Collins, K. (2022). STEM Stories: Fostering STEM persistence for underrepresented minority students attending predominantly White institutions. *Journal of Career Development*, 08948453211073706. http://dx.doi.org/10.1177/08948453211073706

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of mixed methods research*, *1*(2), 112-133. http://dx.doi.org/10.1177/1558689806298224

Jones, C. P. (2002). Confronting institutionalized racism. *Phylon, 50* (1), 7-22. http://dx.doi.org/10.2307/4149999

Jordt, H., Eddy, S. L., Brazil, R., Lau, I., Mann, C., Brownell, S. E., King, K., Freeman, S., & Freeman, S. (2017). Values affirmation intervention reduces achievement gap between underrepresented minority and white students in introductory biology classes. *CBE—Life Sciences Education, 16*(3), ar41. https://doi.org/10.1187/cbe.16-12-0351

Kaplan, S. A., & Garrett, K. E. (2005). The use of logic models by community-based initiatives. *Evaluation and program planning, 28*(2), 167-172. http://dx.doi.org/10.1016/j.evalprogplan.2004.09.002

Keels, M. (2019). *Campus counterspaces: Black and Latinx students' search for community at historically white colleges and universities*. Ithaca: Cornell University Press.

Kendricks, K. D., Nedunuri, K. V., & Arment, A. R. (2013). Minority student perceptions of the impact of mentoring to enhance academic performance in STEM disciplines. *Journal of STEM Education: Innovations and Research*, *14*(2), 38.

Lakin, J. M., & Elliott, D. C. (2016). STEMing the shock: Examining transfer shock and its impact on STEM major and enrollment persistence. *Journal of The First-Year Experience & Students in Transition*, *28*(2), 9-31.

Ladson-Billings, G. (1998). Just what is critical race theory and what's it doing in a nice field like education? *International Journal of Qualitative Studies in Education, 11*(1), 7-24. http://dx.doi.org/10.1080/095183998236863

LaNasa, S. M., Olson, E., & Alleman, N. (2007). The impact of on-campus student growth on first-year student engagement and success. *Research in higher education*, *48*(8), 941-966. http://dx.doi.org/10.1007/s11162-007-9056-5

Lee, M. J., Collins, J. D., Harwood, S. A., Mendenhall, R., & Huntt, M. B. (2020). “If you aren’t White, Asian or Indian, you aren’t an engineer”: racial microaggressions in STEM education. *International Journal of STEM Education*, *7*(1), 1-16. http://dx.doi.org/10.1186/s40594-020-00241-4

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. http://dx.doi.org/10.1006/jvbe.1994.1027

Lent, R. W., Miller, M. J., Smith, P. E., Watford, B. A., Lim, R. H., Hui, K., Morrison, M. A., Wilkins, G., & Williams, K. (2013). Social cognitive predictors of adjustment to engineering majors across gender and race/ethnicity. *Journal of Vocational Behavior*, *83*(1), 22-30. https://doi.org/10.1016/j.jvb.2013.02.006

Lisberg, A., & Woods, B. (2018). Mentorship, mindset and learning strategies: An integrative approach to increasing underrepresented minority student retention in a STEM undergraduate program. *Journal of STEM Education*, *19*(3).

Louis Stoke North Star STEM Alliance. (2016). *North Star STEM Alliance: The internal evaluation plan (phase III).* Unpublished program document.

Louis Stoke North Star STEM Alliance. (2020). *2020 annual report: Key indicators.* Unpublished program document.

Love, A. J. (1991). *Developing systematic internal evaluation*. SAGE Publications, Inc., https://dx.doi.org/10.4135/9781412984546

MacDonald, B. (1974). Evaluation and the control of education. In *Curriculum Evaluation Today: Trends and implications,* 125-136. MacMillan Education.

Magnin, M., Jeannot, E., Voahangy, R., & Stoll, B. (2018). A realistic evaluation approach highlighted the success factors and difficulties of an innovative and comprehensive malnutrition programme in Madagascar. *Acta Paediatrica, 107*(9), 1570-1580. http://dx.doi.org/10.1111/apa.14267

Mahatmya, D., Morrison, J., Jones, R. M., Garner, P. W., Davis, S. N., Manske, J., Berner, N., Johnson, A., & Ditty, J. (2017). Pathways to undergraduate research experiences: A multi-institutional study. *Innovative Higher Education*, *42*(5), 491-504. http://dx.doi.org/10.1007/s10755-017-9401-3

Mansoor, A. K. (2003). *Realist Evaluation in practice: health and social work.* Sage.

Manzano, A. (2016). The craft of interviewing in Realist Evaluation. *Evaluation, 22*(3), 342-360. http://dx.doi.org/10.1177/1356389016638615

Maton, K. I., Domingo, M. R. S., Stolle-McAllister, K. E., Zimmerman, J. L., & Hrabowski III, F. A. (2009). Enhancing the number of African Americans who pursue STEM PhDs: Meyerhoff Scholarship Program outcomes, processes, and individual predictors. *Journal of women and minorities in science and engineering*, *15*(1). http://dx.doi.org/10.1615/JWomenMinorScienEng.v15.i1.20

McGee, E. O., & Martin, D. B. (2011). "You would not believe what I have to go through to prove my intellectual value!" Stereotype management among academically successful Black mathematics and engineering students. *American Educational Research Journal, 48*(6). http://dx.doi.org/10.3102/0002831211423972

Mertens, D. M., & Wilson, A. T. (2018). *Program evaluation theory and practice.* Guilford Publications.

Miller, K., & Servaty-Seib, H. L. (2016). First-year students' loss experiences and institutional belongingness in the transition to college. *Journal of The First-Year Experience & Students in Transition*, *28*(2), 53-72.

Milner, D. I., Horan, J. J., & Tracey, T. J. (2014). Development and evaluation of STEM interest and self-efficacy tests. *Journal of Career Assessment, 22*(4), 642-653. http://dx.doi.org/10.1177/1069072713515427

Mitchell, T. R., Holtom, B. C., Lee, T. W., Sablynski, C. J., & Erez, M. (2001). Why people stay: Using job embeddedness to predict voluntary turnover. *Academy of management journal*, *44*(6), 1102-1121. http://dx.doi.org/10.2307/3069391

Mixson, J. (2015). *North Star STEM Alliance evaluation* [PowerPoint slides]. Unpublished internal document.

Morganson, V. J., Major, D. A., Streets, V. N., Litano, M. L., & Myers, D. P. (2015). Using embeddedness theory to understand and promote persistence in STEM majors. *The Career Development Quarterly*, *63*(4), 348-362. http://dx.doi.org/10.1002/cdq.12033

Musah, R. A., & Ford, M. (2017). Peer-Based Supplemental Instruction in STEM: Differences in Effectiveness Across Transfer and Nontransfer Undergraduates. *Journal of Research on Educational Effectiveness*, *10*(3), 596-618. http://dx.doi.org/10.1080/19345747.2016.1213341

National Center for Educational Statistics. (2018). *The Condition of Education 2018 (NCES 2018-144).* https://nces.ed.gov/pubs2018/2018144.pdf

National Science and Technology Council. (2018, December). *Charting a course for success: America's strategy for STEM education.* https://www.whitehouse.gov/wp-content/uploads/2018/12/STEM-Education-Strategic-Plan-2018.pdf

National Science Board. (2022, January). *U.S. and Global STEM Education and Labor Force.* https://ncses.nsf.gov/pubs/nsb20221/u-s-and-global-stem-education-and-labor-force

National Science Foundation. (n.d.). *Louis Stokes Alliances for Minority Participation*

(LSAMP). https://www.nsf.gov/pubs/2017/nsf17579/nsf17579.htm

National Science Foundation. (2011). *Empowering the nation through discovery and innovation (Fiscal Year 2011–2016).* www.nsf.gov/news/strategicplan/nsfstrategicplan_2011_2016.pdf.

National Science Board (2018). *Science and engineering indicators 2018* (Tech. Rep. No.NSB-2018-1). https://www.nsf.gov/statistics/indicators

National Science Foundation. (2018). *NSF STEM classification of instructional programs (2017-2018 WebAMP*). https://www.ccm.edu/wp-content/uploads/2019/03/NSF-STEM-Classification-of-Instructional-Programs.pdf

Nuñez, A. M., & Yoshimi, J. (2017). A phenomenology of transfer: Students' experiences at a receiving institution. *Innovative Higher Education*, *42*(2), 173-187. http://dx.doi.org/10.1007/s10755-016-9374-7

O'Cathain, A. (2010). Assessing the quality of mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, (531-554). Sage.

Office of Science and Technology Policy. (2019, October). *Progress report on the federal implementation of the STEM education strategic plan*. https://www.whitehouse.gov/wp-content/uploads/2019/10/Progress-Report-on-the-Federal-Implementation-of-the-STEM-Education-Strategic-Plan.pdf

Olimpo, J. T., Fisher, G. R., & DeChenne-Peters, S. E. (2016). Development and evaluation of the Tigriopus course-based undergraduate research experience: Impacts on students' content knowledge, attitudes, and motivation in a major's introductory biology course. *CBE—Life Sciences Education*, *15*(4), ar72. http://dx.doi.org/10.1187/cbe.15-11-0228

Olson, S., & Labov, J. B. (2012). *Community Colleges in the Evolving STEM Education Landscape: Summary of a Summit.* National Academy of Engineering (NAE). https://nap.nationalacademies.org/catalog/13399/community-colleges-in-the-evolving-stem-education-landscape-summary-of

Olson, S., & Riordan, D. G. (2012). Engage to excel: producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. Report to the president. *Executive Office of the President*.

Ong, M., Smith, J. M., & Ko, L. T. (2018). Counterspaces for women of color in STEM higher education: Marginal and central spaces for persistence and success. *Journal of Research in Science Teaching*, *55*(2), 206-245. http://dx.doi.org/10.1002/tea.21417

Ortiz, A. M., & Sriraman, V. (2015). Exploring faculty insights into why undergraduate college students leave STEM fields of study-A three-part organizational self-study. *American Journal of Engineering Education*, *6*(1), 43.

Oseguera, L., Park, H., De Los Rios, M., Aparicio, E., & Johnson, R. (2020). Examining the Role of Scientific Identity in Black Student Retention in a STEM Scholar Program. *The Journal of Negro Education, 88*(3), 229-248. http://dx.doi.org/10.7709/jnegroeducation.88.3.0229

Packard, B. W. L. (2016). *Successful STEM mentoring initiatives for underrepresented students: A research-based guide for faculty and administrators*. Stylus Publishing, LLC.

Palmer, R. T., Maramba, D. C., & Dancy II, T. E. (2011). A qualitative investigation of factors promoting the retention and persistence of students of color in STEM. *Journal of Negro Education*, *80*(4), 491-504.

Patton, M. Q. (1987). *How to use qualitative methods in evaluation (No. 4).* Sage.

Patton, M. Q. (2008). *Utilization-focused evaluation*. Sage publications.

Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice*. Sage publications.

Pawson, R. & Tilley, N. (1997). *Realistic evaluation.* Sage.

Pawson, R., & Tilley, N. (2004). Realist evaluation: The magenta book guidance notes. *London: Cabinet Office. Retrieved May 12*, 2008.

Pawson, R. (2013). *The science of evaluation: a realist manifesto*. Sage.

Pedraza, L., & Chen, R. (2022). Examining motivator factors of STEM undergraduate persistence through two-factor theory. *The Journal of Higher Education*, *93*(4), 532-558. http://dx.doi.org/10.1080/00221546.2021.1999722

Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The journal of educational research*, *96*(1), 3-14. http://dx.doi.org/10.1080/00220670209598786

Perez, T., Cromley, J. G., & Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. *Journal of educational psychology*, *106*(1), 315. http://dx.doi.org/10.1037/a0034027

Pierszalowski, S., Bouwma-Gearhart, J., & Marlow, L. (2021). A systematic review of barriers to accessing undergraduate research for STEM students: Problematizing under-researched factors for students of color. *Social Sciences, 10*(9), 328. http://dx.doi.org/10.3390/socsci10090328

Plano Clark, V. L., & Creswell, J. W. (2014). *Understanding research: A consumer's guide*. Pearson Higher Ed.

Pommier, J., Guével, M. R., & Jourdan, D. (2010). Evaluation of health promotion in schools: a realistic evaluation approach using mixed methods. *BMC public health, 10*(1), 43. http://dx.doi.org/10.1186/1471-2458-10-43

Porter, S., McConnell, T., Clarke, M., Kirkwood, J., Hughes, N., Graham-Wisener, L., Regan, J., McKeown, M., McGrillen, K., & Reid, J. (2017). A critical realist evaluation of a music therapy intervention in palliative care. *BMC palliative care*, *16*(1), 70. http://dx.doi.org/10.1186/s12904-017-0253-5

Potvin, G., & Hazari, Z. (2013). The development and measurement of identity across the physical sciences. In *Proceedings of the 2013 Physics Education Research Conference, Portland, OR*, 281-284. http://dx.doi.org/10.1119/perc.2013.pr.058

QSR International Pty Ltd. (2020). *NVivo (Version 12)* [Computer Software]. Provided at: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

R Core Team (2017). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Reyes, M. E. (2011). Unique challenges for women of color in STEM transferring from community colleges to universities. *Harvard Educational Review, 81*(2), 241-263. http://dx.doi.org/10.17763/haer.81.2.324m5t1535026g76

Riegle-Crumb, C., King, B., & Irizarry, Y. (2019). Does STEM Stand Out? Examining Racial/Ethnic Gaps in Persistence Across Postsecondary Fields. *Educational Researcher, 48*(3), 133–144. http://dx.doi.org/10.3102/0013189X19831006

Robinson-Perez, A., Marzell, M., & Han, W. (2020). Racial microaggressions and psychological distress among undergraduate college students of color: Implications for social work practice. *Clinical Social Work Journal*, *48*(4), 343-350. http://dx.doi.org/10.1007/s10615-019-00711-5

Robnett, R. D., Nelson, P. A., Zurbriggen, E. L., Crosby, F. J., & Chemers, M. M. (2019). The form and function of STEM research mentoring: A mixed-methods analysis focusing on ethnically diverse undergraduates and their mentors. *Emerging Adulthood*, *7*(3), 180-193. http://dx.doi.org/10.1177/2167696818758734

Rock, A. D., & Garavan, T. N. (2006). Reconceptualizing developmental relationships. *Human Resource Development Review, 5*(3), 330-354. http://dx.doi.org/10.1177/1534484306290227

Rogers, P. J., & Weiss, C. H. (2007). Theory-based evaluation: Reflections ten years on: Theory-based evaluation: Past, present, and future. *New directions for evaluation*, *2007*(114), 63-81. http://dx.doi.org/10.1002/ev.225

Rogulkin, D. (2011). *Predicting 6-year graduation and high-achieving and at-risk students*. Fresno State Institutional Research, Assessment & Planning. http://www.fresnostate.edu/academics/oie/documents/documents-research/2011/data%20mining%20report1.pdf

Ross, A., & Willson, V. L. (2017). Independent samples T-test. In *Basic and advanced statistical tests* (13-16). Sense Publishers, Rotterdam. http://dx.doi.org/10.1007/978-94-6351-086-8_3

Rossi, P. H., & Freeman, H. E. (1993). *Evaluation: A systematic approach*. Sage.

Rubin, H. J., & Rubin, I. S. (2011). *Qualitative interviewing: The art of hearing data*.

Sage.

Russell, S. H., Hancock, M. P., & McCullough, J. (2007). Benefits of undergraduate research experiences. *Science*, *316*(5824), 548-549. http://dx.doi.org/10.1126/science.1140384

Saldaña, J. (2021). *The coding manual for qualitative researchers*. Sage.

Schenk, L., Sentse, M., Lenkens, M., Nagelhout, G. E., Engbersen, G., & Severiens, S. (2021). Instrumental Mentoring for Young Adults: A Multi-Method Study. *Journal of Adolescent Research*, *36*(4), 398–424. http://dx.doi.org/10.1177/0743558420979123

Schudde, L. T. (2011). The causal effect of campus residency on college student retention. *The Review of Higher Education*, *34*(4), 581-610. http://dx.doi.org/10.1353/rhe.2011.0023

Schwandt, T. (2015). *Evaluation foundations revisited: Cultivating a life of the mind for practice*. Stanford, California: Stanford Business Books.

Scriven, M. (1991). *Evaluation thesaurus*. Sage Publication.

Scriven, M. (1998). Minimalist theory: The least theory that practice requires. *American Journal of Evaluation*, *19*, 57-70. http://dx.doi.org/10.1016/S1098-2140(99)80180-5

Sens, D. A., Cisek, K. L., Garrett, S. H., Somji, S., Dunlevy, J. R., Sens, M. A., Conway, P., & Doze, V. A. (2017). STEERing an IDeA in undergraduate research at a rural research intensive university. *Academic Pathology*, *4*, 2374289517735092. https://doi.org/10.1177/2374289517735092

Shapiro, D., Dundar, A., Wakhungu, P.K., Yuan, X., Nathan, A, & Hwang, Y. (2016, September). *Time to Degree: A National View of the Time Enrolled and Elapsed for Associate and Bachelor's Degree Earners* (Signature Report No. 11). Herndon, VA: National Student Clearinghouse Research Center.

Short, L., Hennessy, M., & Campbell, J. (1996). Tracking the work. *In Family violence: Building a coordinated community response: A guide for communities*, 59-72.

Simon, R. A., Aulls, M. W., Dedic, H., Hubbard, K., & Hall, N. C. (2015). Exploring student persistence in STEM programs: A motivational model. *Canadian Journal of Education*, *38*(1), 1.

Smith, K. C., Poleacovschi, C., Feinstein, S., & Luster-Teasley, S. (2022). Ethnicity, Race, and Gender in Engineering Education: The Nuanced Experiences of Male and Female Latinx Engineering Undergraduates Targeted by Microaggressions. *Psychological Reports*, *0*(0). http://dx.doi.org/10.1177/00332941221075766

Solórzano, D. G., & Yosso, T. J. (2001). Critical race and LatCrit theory and method: Counter-storytelling. *International journal of qualitative studies in education*, *14*(4), 471-495. http://dx.doi.org/10.1080/09518390110063365

Solórzano, D., Villalpando, O., Torres, C., & Mitchell, T. (1998). Critical race theory, marginality, and the experience of minority students in higher education. *Emerging issues in the sociology of education: Comparative perspectives*, 211-224.

Soria, K. M. (2014, March). *LSAMP North Star Evaluation Overview* [PowerPoint slides]. Unpublished internal document.

Sridharan, S., & Nakaima, A. (2012). Towards an evidence base of theory-driven evaluations: Some questions for proponents of theory-driven evaluation. *Evaluation, 18*(3), 378-395. http://dx.doi.org/10.1177/1356389012453289

Stolle-McAllister, K., & Carrillo, A. (2011). The Meyerhoff way: How the Meyerhoff scholarship program helps black students succeed in the sciences. *Journal of science education and technology*, *20*(1), 5-16. http://dx.doi.org/10.1007/s10956-010-9228-5

Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, & applications*. San Francisco, CA: Jossey-Bass.

Sue, D. W., Capodilupo, C. M., & Holder, A. M. B. (2008). Racial microaggressions in the life experience of Black Americans. *Professional Psychology: Research and Practice, 39*(3), 329–336. http://dx.doi.org/10.1037/0735-7028.39.3.329

Sweeder, R., Kursav, M., & Valles, S. (2021). A Cohort Scholarship Program That Reduces Inequities in STEM Retention. *Journal of STEM Education, 22*(1), 5-13.

Syed, M., Zurbriggen, E. L., Chemers, M. M., Goza, B. K., Bearman, S., Crosby, F. J., Shaw, J. M., Hunter, L., & Morgan, E. M. (2019). The role of self-efficacy and identity in mediating the effects of STEM support experiences. *Analyses of Social Issues and Public Policy*, *19*(1), 7-49. https://doi.org/10.1111/asap.12170

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches* (Vol. 46). Sage.

Tashakkori, A., & Teddlie, C. (2021). *Sage handbook of mixed methods in social & behavioral research*. SAGE publications. http://dx.doi.org/10.4135/9781506335193

Thomas, D. A. (2001). The truth about mentoring minorities. Race matters. *Harvard* *business review*, *79*(4), 98-107.

Tilley, N. (2000, September). Realistic evaluation: an overview. In *founding conference of the Danish Evaluation Society (Vol. 8).*

Tinto, V. (1975). Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research*, *45*(1), 89–125. https://doi.org/10.3102/00346543045001089

Tinto, V. (1997). Classrooms as communities: exploring the educational character of student persistence. *The Journal of Higher Education*, *68*(6), 599-623. http://dx.doi.org/10.2307/2959965

Torres-Harding, S. R., Andrade, A. L., & Romero Diaz, C. E. (2012). The racial microaggressions scale (RMAS): A new scale to measure experiences of racial microaggressions in people of color. *Cultural Diversity and Ethnic Minority Psychology*, *18*(2), 153–164. http://dx.doi.org/10.1037/a0027658

Tupper, D. H., Leitherer, B., Sorkin, S., & Gore, M. E. (2010). Strategies for Increasing IT Enrollment: Recruiting, Retaining and Encouraging the Transfer of Women and Underrepresented Groups to Four-Year Colleges. *Information Systems Education Journal*, *8*(54), n54.

Van der Knaap, L. M., Leeuw, F. L., Bogaerts, S., & Nijssen, L. T. (2008). Combining Campbell standards and the Realist Evaluation approach: the best of two worlds?. *American journal of evaluation*, *29*(1), 48-57. http://dx.doi.org/10.1177/1098214007313024

Van Dort, L. (2018). *Invisibility Perpetuated: The Complex Economics of Asian*

Minnesotans. Unpublished Manuscript. https://caalmn.org/wpcontent/uploads/2021/03/Invisibility-Perpetuated-the-Complex-Economics-of-Asian-Minnesotans.pdf

VanVoorhis, C. W., & Morgan, B. L. (2007). Understanding power and rules of thumb for determining sample sizes. *Tutorials in quantitative methods for psychology*, *3*(2), 43-50.

Volkov, B. B. (2011). Internal evaluation a quarter-century later: A conversation with Arnold J. Love. In B. B. Volkov & M. E. Baron (Eds.), *Internal evaluation in the 21st century*. *New Directions for Evaluation, 132*, 5–12. http://dx.doi.org/10.1002/ev.392

Wang, X. (2016). Course-taking patterns of community college students beginning in STEM: Using data mining techniques to reveal viable STEM transfer pathways. *Research in Higher Education*, *57*(5), 544-569. http://dx.doi.org/10.1007/s11162-015-9397-4

Wang, X., Lee, S. Y., & Prevost, A. (2017). The role of aspirational experiences and behaviors in cultivating momentum for transfer access in STEM: Variations across gender and race. *Community College Review*, *45*(4), 311-330. http://dx.doi.org/10.1177/0091552117724511

Weiss, C. H. (1988). Reports on topic areas: Evaluation for decisions: Is anybody there? Does anybody care?. *Evaluation practice*, *9*(1), 5-19. http://dx.doi.org/10.1177/109821408800900101

Weiss, C. H. (1997). Theory-based evaluation: past, present, and future. *New directions for evaluation*, *76*, 41-55. http://dx.doi.org/10.1002/ev.1086

Westhorp, G. (2012). Using complexity-consistent theory for evaluating complex systems. *Evaluation, 18*(4), 405-420. http://dx.doi.org/10.1177/1356389012460963

Westhorp, G. (2013). Developing complexity-consistent theory in a realist investigation. *Evaluation, 19*(4), 364-382. http://dx.doi.org/10.1177/1356389013505042

Westhorp, G. (2014). Realist impact evaluation: an introduction. *London: Overseas Development Institute*, 1-12.

Whitcomb, K. M. & Singh, C. (2021). Underrepresented minority students receive lower grades and have higher rates of attrition across STEM disciplines: A sign of inequity?, *International Journal of Science Education, 43*(7), 1054-1089. http://doi.org/10.1080/09500693.2021.1900623

White, H. (2009). Theory-based impact evaluation: principles and practice. *Journal of development effectiveness*, *1*(3), 271-284. http://dx.doi.org/10.1080/19439340903114628

Wilson, Z. S., Holmes, L., Degravelles, K., Sylvain, M. R., Batiste, L., Johnson, M., McGuire, S. Y., Pang, S. S., & Warner, I. M. (2012). Hierarchical mentoring: A transformative strategy for improving diversity and retention in undergraduate STEM disciplines. *Journal of Science Education and Technology, 21*(1), 148-156. http://dx.doi.org/10.1007/s10956-011-9292-5

Wilson, Z., Iyengar, S., Pang, S., Warner, I., & Luces, C. (2012). Increasing Access for Economically Disadvantaged Students: The NSF/CSEM & S-STEM Programs at Louisiana State University. *Journal of Science Education and Technology, 21*(5), 581-587. http://dx.doi.org/10.1007/s10956-011-9348-6

Woolcock, M. (2009). Toward a plurality of methods in project evaluation: a contextualized approach to understanding impact trajectories and efficacy. *Journal of Development Effectiveness 1*, 1–14. http://dx.doi.org/10.1080/19439340902727719

Wright, B., Lee, S., Zaromatidis, K., & Huk, O. (2021). Increasing Retention of STEM Students with Financial Need via the NSF S-STEM DESIRE Scholarship Program. *Journal of STEM Education, 22*(4), 24-38.

Xu, Y. J. (2018). The Experience and Persistence of College Students in STEM Majors. *Journal of College Student Retention: Research, Theory & Practice, 19*(4), 413-432. http://dx.doi.org/10.1177/1521025116638344

Xu, Y. J., & Webber, K. L. (2016). College student retention on a racially diverse campus: A theoretically guided reality check. *Journal of College Student Retention: Research, Theory & Practice, 0*(0), 1-27. http://dx.doi.org/10.1177/1521025116643325

Yosso, T. J. (2013). *Critical race counterstories along the Chicana/Chicano educational pipeline*. Routledge. http://dx.doi.org/10.4324/9780203624821

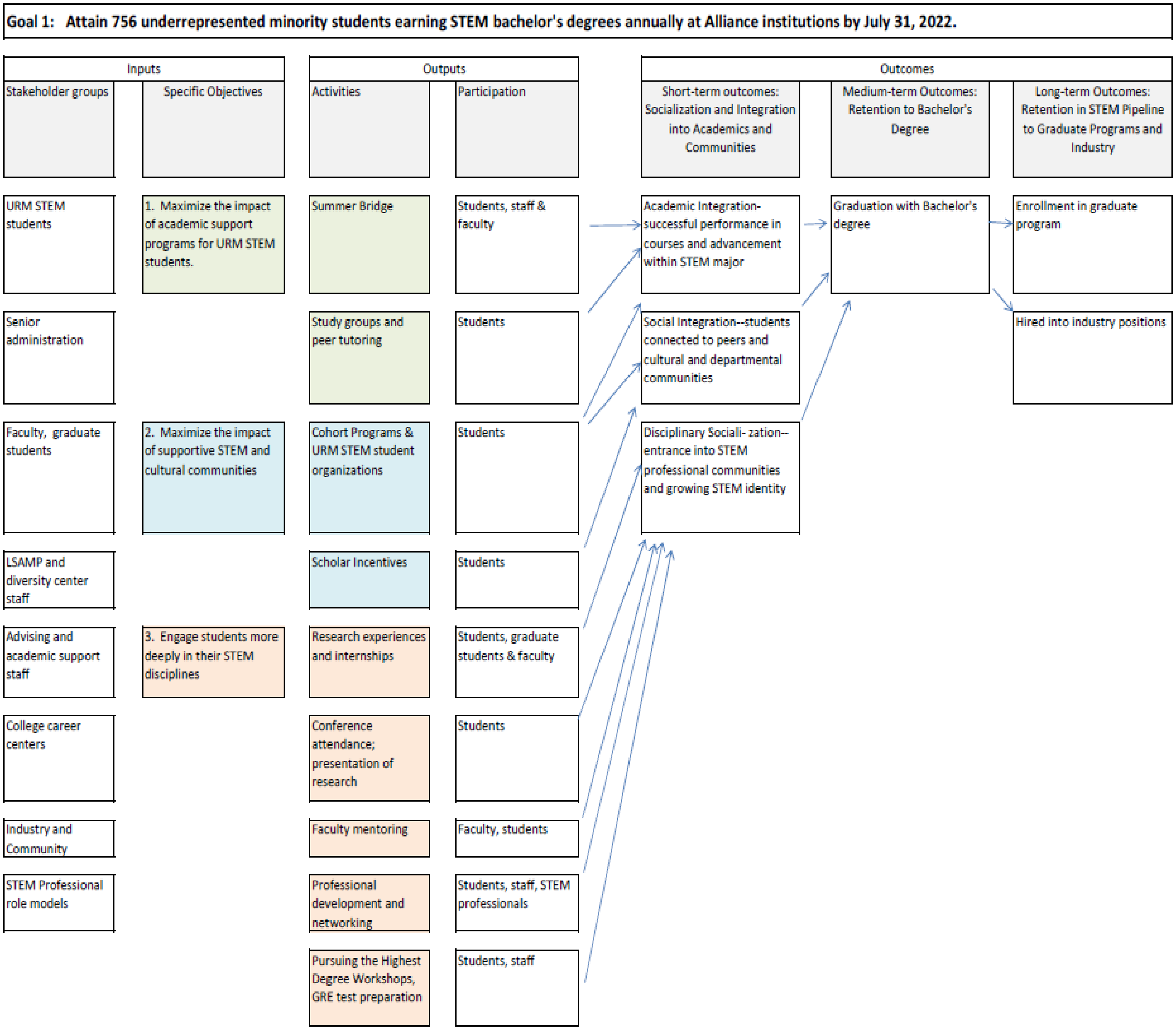
Yosso, T., Smith, W., Ceja, M., & Solórzano, D. (2009). Critical race theory, racial microaggressions, and campus racial climate for Latina/o undergraduates. *Harvard Educational Review, 79*(4), 659-691. http://dx.doi.org/10.17763/haer.79.4.m6867014157m707l

Zaniewski, A. M., & Reinholz, D. (2016). Increasing STEM success: a near-peer mentoring program in the physical sciences. *International Journal of STEM Education, 3*(1), 1-12. http://dx.doi.org/10.1186/s40594-016-0043-2

Zoom Video Communications Inc. (2022). *Zoom (Version 5.12.3)* [Computer software].

Appendix A

North Star STEM Alliance: The Logic Model (Phase III)



Goal 2: Deepen Alliance collaborations to improve students' pathways through all stages of transfer from community colleges to degrees at 4-year institutions.

INPUTS		OUTPUTS		OUTCOMES		
Stakeholder Groups	Specific Objectives	Activities	Participation	Short	Medium	Long
Minority STEM target students, including prospective transfer students	1. Increase twofold the proportion of transfer students that are retained in STEM after three years' enrollment in Alliance 4-year institutions.	Activities referenced in Goal 1 related to Academic Integration, Social Integration and Discipline Socialization	URM STEM students, faculty, and student services staff	Faster adjustment to campus, academic rigor, major, social community	Graduation with Bachelor's Degree	Retention in STEM Pipeline; enrollment in STEM graduate programs, employment in industry
Alliance Leadership Team, site coordinators and faculty mentors	2. Identify performance metrics and institutional practices across Alliance that inform and influence the success of transfer students.	Transfer Stakeholder Group meets four times over two years to share knowledge of transfer processes, prototype new retention strategies and define new metrics.	North Star site coordinators, campus colleagues, including transfer admissions counselors, faculty mentors, academic advisors, diversity specialists	Clarified transfer processes across Alliance; Refined retention strategies and associated metrics.	Greater retention of transfers in STEM majors	More transfer students graduate with STEM degrees
Campus student services staff: Transfer admissions counselors, academic advisors, diversity specialists	3. Engage students and administrators at community colleges beyond the Alliance to build stronger STEM transfer pathways to Alliance institutions.	Transfer Program activities, including campus visits, transfer admissions advising, opportunities to meet faculty and STEM professionals, peer mentoring.	North Star site coordinators, transfer admissions counselors, faculty mentors, academic advisors, diversity specialists; prospective transfer students	Strudents have more positive transfer experience: more initial connections to faculty, staff & students; more sense of belonging; stronger performance in courses.	More students retained in STEM majors	More transfer students graduate in STEM with less time to degree

Goal 3: Foster institutional change toward greater diversity and inclusion on Alliance campuses.

Activities vary across institutional partners

Inputs		Outputs		Outcomes		
Stakeholder Groups	Specific Objectives	Activities	Participation	Short	Medium	Long
Senior administrators	Enhance the academic and collegiate climate for success of underrepresented students in STEM through faculty engagement and institutional change.	Campus Climate Initiatives. Examples include campus climate surveys, inclusive excellence messaging, bias response and referral networks	Whole campus community	More welcoming, diverse, and inclusive campus. Minority students' greater sense of belonging	Greater retention of URM students in STEM majors	More students graduating with STEM degrees;
Faculty, student services and Equity/Diversity staff		Implicit bias and cultural diversity training initiatives. Examples include cultural competency certificate programs, implicit bias workshops, mentor training	Faculty, staff, students, diversity trainers	More inclusive and welcoming classrooms and services		
Students		Culturally responsive pedagogy, incorporating multicultural perspectives and cultural knowledge into the curriculum	Faculty, elders from cultural communities	Minority STEM students more engaged in courses; greater sense of belonging		
Elders from cultural communities		Active learning strategies incorporated into lecture courses	Faculty	Courses more open to varied learning styles		
Institutional Diversity plans		Recruit and retain STEM faculty of color	Administrators and faculty	Minority STEM students more engaged in courses; stronger cultural identity, greater sense of belonging		

Appendix B

North Star STEM Alliance: The Internal Evaluation Plan (Phase III)

Goal 1: Attain 756 underrepresented minority students earning STEM bachelor's degrees annually at Alliance institutions by July 31, 2022.				
Objective	*Activities*	*Outcomes*	*Quantitative measures*	*Qualitative measures*
Objective 1a: Maximize the impact of academic support programs for URM STEM students.	Faculty mentoring, group study, supplemental instruction, tutoring, summer online courses in mathematics	Academic integration – Students perform successfully in courses and advance within STEM majors	1) Target audience analysis correlating to Pell eligibility and first-generation status 2) Engagement in retention and co-curricular programs, analyzed by race/ethnicity, gender, socioeconomic factors	Potential qualitative evaluation studies informed by interpretive research component assessing broad experience of URM STEM students.
Objective 1b: Maximize the impact of supportive STEM and cultural communities through social networking, peer mentorship, and student organizations.	Alliance Kick-Off; Scholar engagement and cohort programs; attendance at national conferences; engagement in cultural STEM student organizations	Social Integration--Students connected to peers and cultural and departmental communities	1) Engagement in retention and co-curricular programs and cultural STEM student organizations, analyzed by race/ethnicity, gender, socioeconomic factors	Potential qualitative evaluation studies informed by interpretive research component assessing broad experience of URM STEM students.
Objective 1c: Engage students more deeply in their STEM disciplines through research opportunities, faculty mentorship, and preparation for graduate school.	Student placement in research within Alliance, REUs, international research; Summer Opportunities Preview; Research and Native Knowledge workshop; Pursuing the Highest Degree workshop; GRE Test Preparation	Disciplinary Socialization—Students enter into STEM professional communities and grow their identity as a STEM professional	1) Number of students placed in local, national and international research opportunities* 2) Faculty engaged in mentoring* 3) Attendance at graduate school preparation, test prep, and Native research workshop	1) Student reflections following research experiences 2) Event evaluations following Native workshop, PhD workshop. 3) Potential qualitative studies informed by interpretive research component
General Goal 1 Degree Attainment	Activities described above.	Attain 756 underrepresented minority students earning STEM bachelor's degrees annually at Alliance institutions by July 31, 2022.	1) STEM Bachelor's degrees by race/ethnicity earned since Alliance formation ** 2) Enrollment by discipline; enrollment by race/ethnicity ** 3) 4-year, 5-year, and 6-year graduation rates	
Goal 2: Deepen Alliance collaborations to improve students' pathways through all stages of transfer from community colleges to degrees at 4-year institutions.				
Objective	*Activities*	*Outcomes*	*Quantitative measures*	*Qualitative measures*

Objective 2a) Increase twofold the proportion of transfer students that are retained in STEM after three years' enrollment in Alliance 4-year institutions.	Scholar engagement programs, social gatherings, cohorts, peer mentors within departments or related majors, First Year experience courses for transfers Interventions prototyped by Transfer Stakeholder Group	Double the proportion of transfer students are retained or graduated in STEM three years after the Fall 2019 cohort enrolls, compared to Fall 2012 cohort.	1) Outcomes of Minority Transfer Students into 4-Year Alliance Institutions[1]—Status by Cohort Entry Term	1) Survey or interviews with transfer students at end of their first year, or alternatively, assess the interventions developed in Transfer Stakeholder Group
Objective 2b) As an Alliance, identify performance metrics and institutional practices within and across colleges and universities that inform and influence the success of transfer students.	Transfer Stakeholder Group identifies new institutional practices for prototyping and implementation. Website aids to navigating transfer process	1) Refined strategies and metrics on preparing students for transfer, retaining new transfer students, and shared knowledge of STEM pathways and transfer processes across Alliance institutions.	1) Metrics to be identified by Transfer Stakeholder Group and carried out by UMTC and partner institutions which accept transfers. 2) Transfer engagement analyzed within study in 1a) of Engagement in retention and co-curricular programs.	1) Qualitative studies to be identified by Transfer Stakeholder Group. 2) Survey on navigating transfer pathways by transfer students before matriculation to their 4-year institution
Objective 2c) Engage students and administrators at community colleges beyond the Alliance to build stronger STEM transfer pathways to Alliance institutions.	Transfer Program activities, including campus visits to 4-year institutions, informal Q&A with transfer admissions counselors and academic advisors, networking with student leaders of cultural STEM organizations	Streamlined transfer pathways, better and more immediate academic and social adjustment to the 4-year institution.	1) Counts of students from 2-year institutions outside of Alliance served by prospective transfer programming. * 2) Recruitment activity and demographics served. *	
Goal 3: Foster institutional change toward greater diversity and inclusion on Alliance campuses.				
Objective	*Activities*	*Outcomes*	*Quantitative measures*	*Qualitative measures*
Objective 3a) Enhance the academic and collegiate climate for success of underrepresented students in STEM through faculty engagement and institutional change.	Campus climate surveys, Implementation of **cultural competency certificate programs.** Recruiting and retaining faculty of color Partner institutions create system frameworks for co-curricular programs & services.	Students' improved sense of belonging, identity with community and institution, comfort within institutional climate, engagement Program maps of	1) Counts of gains and losses in retaining faculty of color***	1) New initiatives and progress reported by Governing Board members ***

Note. Studies carried out by graduate research assistants with quantitative data compiled by institutional researchers and analysts, unless data reported as indicated: *site coordinators; **WebAMP data; *** Governing Board members or staff supporting Governing Board members.

Appendix C

The Correlation Matrix of Predictors and Covariates Involved in the Process of Analysis

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Graduation Status	1.000													
2. First-generation Status	*** -0.065	1.000												
3. Male	** -0.022	*** -0.024	1.000											
4. Age	* 0.013	*** 0.167	*** 0.081	1.000										
5. Total Transferred Credits	*** 0.117	*** 0.076	*** 0.069	*** 0.586	1.000									
6. First-Semester Credits Pass Rate	*** 0.286	*** -0.100	*** -0.035	*** -0.131	*** -0.03	1.000								
7. First-Semester GPA	*** 0.337	*** -0.147	*** -0.063	*** -0.104	*** 0.044	*** 0.666	1.000							
8. Greater state area	*** -0.038	*** 0.083	0.006	*** -0.028	*** -0.023	** -0.019	*** -0.051	1.000						
9. Reciprocity States	*** 0.025	*** -0.034	*** -0.043	*** -0.059	*** -0.066	*** 0.038	*** 0.044	*** -0.238	1.000					
10. Other States	0.003	*** -0.096	*** -0.024	*** -0.053	*** 0.018	*** 0.037	*** 0.055	*** -0.206	*** -0.132	1.000				
11. Housing Status	*** 0.087	*** -0.229	*** -0.051	*** -0.313	*** -0.182	*** 0.161	*** 0.249	*** -0.166	*** 0.133	*** 0.231	1.000			
12. Pell Grant Eligibility	*** -0.051	*** 0.301	-0.009	*** 0.212	*** 0.112	*** -0.108	*** -0.127	*** 0.055	*** -0.036	*** -0.076	*** -0.210	1.000		
13. Full-time Status	*** 0.040	*** -0.055	-0.007	*** -0.301	*** -0.196	*** 0.072	*** 0.075	0.008	*** 0.025	*** 0.030	*** 0.108	-0.014	1.000	
14. Stopout Years	*** -0.895	*** 0.093	* 0.011	*** 0.027	*** -0.086	*** -0.358	*** -0.424	*** 0.060	*** 0.036	-0.008	*** -0.130	*** -0.690	*** -0.056	1.000

Note. Statistical Significance: * p<.05, ** p<.01, ***p<.001

Appendix D

Research Mentorship Programs Interview Protocol

Introduction and Purpose

Welcome and thank you for being here today. The interview aims to learn about your experience in the research mentorship program, as well as your opinions on related issues. Your sharing is unique and valued by LS-NSSA. Your input will help us understand participants' perspectives on program features. It will inform LS-NSSA's program development to better serve our students. I will lead the conversation by asking questions, but my role is mainly to support you in expressing your opinions and experiences. Please let me know if you have any questions before we start the conversation.

Consent and Confidentiality Disclaimer

- *Confidentiality of information*: Your personal information will not be released in any format. All the direct quotes in any report will be deidentified and coded. This discussion will be recorded for narrative analysis. Only the researcher would have access to these materials. If you feel uncomfortable showing your face, please feel free to turn off your camera anytime during the conversation.
- *The willingness to participate*: Your participation is entirely voluntary. You may end the interview at any time. Your choices and participation will not affect your affiliation with the LS-NSSA program in any way.
- *Comfortableness in sharing*: There is no right or wrong answer, and no judgment will be attached to your opinions, thoughts, and experiences. It is always fine to skip any questions that do not make sense to you.

- *The length of your participation*: Our conversation will last about 30-40 minutes today.

Opening Question

- What did you expect to obtain from the research/mentorship?
  - Do you think you have achieved your initial expectations through the research/mentoring program?

Research and Mentoring Experience

- Who was you mentor? How well do you know about him or her?
- How do you describe the mentorship's influence on your learning experience?
  - What has helped you most from the mentorship?
  - What do you think the mentor could have done better to improve your experience?
- What challenge(s) have you experienced during the program? How did you feel?
  - What did your mentor do to support you and help you get through the challenge(s)?
  - Besides academic support, did you receive emotional support from your mentor?
- How would you perceive/describe the similarity between you and your mentor (In terms of ethnicity, gender, personal experience, goals, etc.)?
  - would you feel more comfortable or helpful to have a mentor that has a higher level of demographic similarity?
- What mentor qualities would you like to see from your future mentors, if there could be any?

Diversity, Equity, and Inclusion

- What do diversity, equity, and inclusion mean to you?
- In general, how would you describe on-campus diversity, equity, and inclusion that have influenced your learning experience in STEM?
  - Do you feel that you have equal access to the resources that you need?
  - If any, could you please describe the moment when you were challenged by a situation where others were behaving in a culturally inappropriate way (e.g., being racially offensive, unequal, exclusive, harmful, disrespectful, etc.)?
  - How would you compare the atmosphere in this research & mentorship program with the general class/lab environment? What makes a difference?
- Being a STEM student from a rich cultural background, can you comment and share your insights on the research/mentoring program from that perspective?
  - How did it feel interacting with co-workers?
  - Did you feel belonged? Did you feel comfortable? Did you feel empowered?
  - How do you describe your role in the workspace?

Program Influence on STEM Education

- How does this research/mentorship program influence your major direction, career choice, or next steps in the STEM discipline?
  - Do you feel more confident to pursue what you expect to achieve?
  - How does it help you shape your science identity?

- What plan do you have in the next semester/year to enhance your personal development in STEM?

Ending Question

- Do you have any other comments on this program or LS-NSSA if I have not covered yet?