UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

Evaluating Supports for Autistic Populations in Postsecondary Education: A Synthesis of Extant Literature and a Single-Case Experimental Design

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirement of the

Degree of

DOCTOR OF PHILOSOPHY

By

Ryan E. Beaty

Norman, Oklahoma

2025

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/341554

Evaluating Supports for Autistic Populations in Postsecondary Education:

A Synthesis of Extant Literature and a Single-Case Experimental Design

A DISSERTATION APPROVED FOR THE

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Corey Peltier, Chair

Dr. Kendra Williams-Diehm

Dr. Malarie Deardorff

Dr. Michael Crowson

Dr. Meg Myers Morgan

© Copyright by Ryan E. Beaty

All Rights Reserved.

Dedication

For my favorite…

Acknowledgements

I'm a man of few words and fewer sentiments. While I can reflect on research for days, turning that lens inward is another matter. It requires a kind of emotional spelunking that I find both foreign and uncomfortable. Nevertheless, the gratitude I feel for the following people demands the effort. Writing this section may prove the most difficult of my entire dissertation…

For Elle. In the quiet, isolating years of this work, your love was the one constant. You are my favorite person, a truth that has grounded me every single day. I don't know that I would have ever started this academic journey without your belief in me, and I know for certain I would not have finished it without your steadfast support. You have sacrificed and celebrated alongside me, even when I hated to celebrate myself, and this accomplishment is as much yours as it is mine. My love for you transcends words, but I will spend the rest of our lives showing you its depth. Thank you for everything. I love you.

To my advisor, Dr. Corey Peltier, who has graciously endured my presence since entering my master's program. Thank you for introducing me to the worlds of single-case and meta-analysis, for instilling in me the importance of transparent, open science in academia, and for your seemingly infinite patience, both with my constant barrage of questions and my habit of answering your questions whenever the class fell silent. Your guidance has been invaluable, and I hope this dissertation marks not the end but the beginning of future collaborations.

To Dr. Kendra Williams-Diehm, I owe a special debt of gratitude. When I was considering graduate school, you did more than offer advice by providing the funds that turned an uncertainty into a reality. Without your belief in me at that critical point, I doubt I ever would have started this journey. Thank you for not only supporting me through two advanced degrees but for personally teaching me the invaluable skill of grant writing and affording me the chance to participate in the process. You gave me a career trajectory, and for that, I will always be thankful.

To Dr. Malarie Deardorff, who accomplished the impossible. Through mere discussions during my first year as a TRAIN Scholar, you made me appreciate qualitative research. I will confess that I still hate the process of thematic analysis, but thanks to your perspective, I can no longer deny its power in policy and advocacy. You taught this stubborn, quantitative-minded researcher that audiences tend to prefer the human element over mere discussion of data. For that invaluable lesson, for your support across my entire graduate career, and for never closing your door on the student perpetually lingering in your hallway, I am incredibly grateful, even though I was not able to secure interviews at the culmination of this study…

To Dr. Michael Crowson, thank you for making quantitative analysis accessible and exciting. I learned an incredible amount through your asynchronous courses, and I kindly beg you to never shutter your YouTube page for the sake of all future graduate students. I deeply appreciated your thoughtful feedback and your thorough use of exemplars, which were undeniably beneficial. Also, thank you for deeming my prospectus comprehensible, which was a shot of confidence I desperately needed going into my dissertation.

To Dr. Meg Myers Morgan, thank you for welcoming an education student into the political science space. Making the commute to Tulsa was worth every minute to learn from your practical, hands-on approach, especially with the population we were able to serve. I am also deeply grateful for your grace as I fumbled my way through facilitating a mock focus group. You gave me a skillset in program evaluation that I applied immediately in my own scholarship and service, and for that, I am incredibly thankful.

To Peighton and Isela, it has been an honor to navigate two advanced degrees alongside you, and I have treasured the friendship and camaraderie we built. While our paths may diverge for now, I am hopeful we will have the opportunity to collaborate again in the future. Thank you for your patience, your support, and for being there through every challenge and success along the way.

To the ZI Crew, I treasured every opportunity we had to work together, and I learned a significant amount from each of you. Whether you are nearing the end or just beginning your own journeys, I know you will all achieve great things. I wish you the best of luck as you reach the finish line. At the very least, you can now feel comfort that the office yeti, who has behaved as a grumpy octogenarian since birth, is no longer trying to freeze you into hibernation.

To those who will never read this, thank you for the lessons taught through your work. To Hidetaka Miyazaki, you showed me the value of perseverance; every academic obstacle was a boss battle, replete with that same rush of accomplishment upon victory. To Eminem, your passion and intricate lyrics were the fuel for my writing sessions. I hope to have a career with the same longevity and impact. Thank you both for the inspiration.

To the unknown readers of this dissertation, thank you for considering my work noteworthy, especially if you're here for these weird acknowledgements. If you stumbled upon this in your own research, I thank you for being an advocate for the Autistic Community and hope this text proves useful in some capacity. If you're reading this for any other reason, well, I can only assume you're a masochist. In my doctoral studies I've learned not to "yuck someone else's yum," an adage I still don't fully understand, so I sincerely hope you're having a good time.

Well, so much for brevity, eh?

Table of Contents

Dedication iv

Acknowledgements v

Table of Contents viii

List of Tables xi

List of Figures xii

Abstract xiii

Chapter 1: Problem Statement 1

Systematic Literature Review 3

Meta-Analysis 4

Single-Case Experimental Design 5

Chapter 2: Literature Review 6

Prior Research 8

Current Review 11

Method 12

Review Methods 12

Search Strategy 12

Inclusion/Exclusion Criteria 14

Screening Procedure 15

Interobserver Agreement on Initial Screening 17

Quality Assessment 18

Study Characteristic Coding 24

Results 24

Study Characteristics 26

Discussion 30

Limitations 33

Implications for Practice and Future Research 33

Chapter 3: Meta-Analysis 35

Method 37

Review Methods and Search Strategy 38

Inclusion/Exclusion Criteria 38

Variable Coding 40

Data Extraction and Estimation of Effects 42

Results 44

Summary of Effects 44
Publication Bias 44
Moderator Analyses 45
Discussion 53
Limitations 55
Implications for Practice and Future Research 55
Chapter 4: Single-Case Experimental Design 58
Study Purpose 59
Methods **60**
Participants 60
Setting 63
Materials 63
Experimental Design and Procedures 68
Treatment Fidelity 74
Preliminary Data Analyses 75
Study Deviations 79
Results **80**
Chase 80
Alyx 92
Discussion 107
Extant Literature Connection 112
Contextualizing the Effect Size 114
Methodological Considerations, Strengths, and Limitations 115
Implications for Practice and Future Research 117
Chapter 5: Discussion and Future Research 119
Synthesized Discussion of Findings 119
Importance of Null or Negative Findings 121
Limitations 122
Future Research 124
References **128**
Appendix A: Literature Review and Meta-Analysis (Tables and Figures) **143**
Table 1. Research Design 143
Table 2. Participant Demographics 147
Table 3. Findings and Social Validity 148
Table 4. NTACT Quality Indicators for Single-Case Experimental Designs 151

Figure 1. Interobserver Agreement (NTACT:C Quality Indicators for SCEDs) 152
Figure 2. Funnel Plot (Effects Level) 153
Appendix B: Single-Case Experimental Design (Tables and Figures) **154**
Figure 1. m-Path Daily Prompts (Screenshot) 154
Figure 2. Fidelity Checklist (Chase [Top]; Alyx [Bottom]) 155
Figure 3. Chase's Belonging (Daily Set) 156
Figure 4. Chase's Positivity (Daily Set) 157
Figure 5. Alyx's Belonging (Daily Set) 158
Figure 6. Alyx's Positivity (Daily Set) 159

List of Tables

Table 1. Postsecondary Enrollment By Disability Classification 7
Table 2. Interobserver Agreement (Article Pool) 18
Table 3. Interobserver Agreement (Quality Assessment By Indicator) 24
Table 4. Multivariate Moderator Analyses 47
Table 5. Participant Demographics 63
Table 6. Chase's Conservative Dual-Criterion (Belonging) 83
Table 7. Chase's Belonging (NAP; LRR) 84
Table 8. Chase's Social Partner Report (Belonging) 84
Table 9. Chase's Conservative Dual-Criterion (Positivity) 88
Table 10. Chase's Positivity (NAP; LRR) 89
Table 11. Chase's Social Partner Report (Positivity) 90
Table 12. Chase's Pre-Post (Revised Social Connectedness Scale–Modified) 91
Table 13. Chase Pre-Post Test of Young Adult Social Skills Knowledge 91
Table 14. Chase Pre-Post (Children's Intervention Rating Profile–Modified) 92
Table 15. Alyx's Conservative Dual-Criterion (Belonging) 96
Table 16. Alyx's Belonging (NAP; LRR) 97
Table 17. Alyx's Social Partner Report (Belonging) 98
Table 18. Alyx's Conservative Dual-Criterion (Positivity) 102
Table 19. Alyx's Positivity (NAP; LRR) 103
Table 20. Alyx's Social Partner Report (Positivity) 104
Table 21. Alyx Pre-Post (Revised Social Connectedness Scale–Modified) 105
Table 22. Alyx Pre-Post Test of Young Adult Social Skills Knowledge 106
Table 23. Alyx Pre-Post (Children's Intervention Rating Profile – Modified) 107

List of Figures

Figure 1. Screening Diagram 16
Figure 2. Funnel Plot of Effects 45
Figure 3. Chase's Belonging (Weekly Mean) 81
Figure 4. Chase's Belonging (Weekly Median) 82
Figure 5. Chase's Positivity (Weekly Mean) 86
Figure 6. Chase's Positivity (Weekly Median) 87
Figure 7. Alyx's Belonging (Weekly Mean) 93
Figure 8. Alyx's Belonging (Weekly Median) 95
Figure 9. Alyx's Positivity (Weekly Mean) 99
Figure 10. Alyx's Positivity (Weekly Median) 101
Figure 11. Forest Plot w/Current Study 115

Abstract

Autistic individuals are enrolling in postsecondary education at increasing rates but continue to experience lower completion rates than their peers, often experiencing challenges including but not limited to social engagement, mental well-being, and sense of belonging. This dissertation includes three manuscripts. The first manuscript was a PRISMA-aligned, systematic review of 15 single-case experimental design (SCED) studies implementing interventions that targeted social, academic, and adaptive behaviors for autistic students enrolled in postsecondary education. Generally, positive outcomes were observed; however, the review revealed significant methodological variability and limitations across the literature, including inconsistent reporting and limited generalizability. Informed by the review, the second manuscript was a meta-analysis of 14 SCED studies to analyze the average effect and explore the heterogeneity. Results indicated a moderate-to-large average effect, albeit with significant heterogeneity. Moderator analyses revealed that intervention effects were stronger under certain, with evidence of potential publication bias identified in the analysis. Results of the meta-analysis guided the development of the third manuscript. A multiple baseline across participants design was implemented to evaluate the Program for the Education and Enrichment of Relational Skills for Young Adults (PEERS®-YA) intervention, using experience sampling methodology (ESM) to collect highly ecological daily data to assess the intervention's effect on participants' sense of belonging and overall positivity. Findings were inconclusive, with visual and statistical analyses not demonstrating a functional relation. However, both participants showed gains on pre-post measures of campus connectedness and reported high social validity for the program. Collectively, these studies indicate that while interventions for autistic undergraduates show promise, their effectiveness is variable. The dissertation underscores the importance of methodological rigor, the need for individualized supports, and the critical value of reporting null or inconclusive findings to build a more balanced and authentic evidence base.

Chapter 1: Problem Statement

Autism diagnoses have steadily risen over the past decade, resulting in more autistic individuals transitioning into adulthood and navigating postsecondary education environments (Dietz et al., 2020; Maenner et al., 2023). The Centers for Disease Control and Prevention (CDC) estimated that approximately 5.5 million autistic adults lived in the United States in 2017, with recent data showing an increased prevalence of 1 in 36 among 8-year-old children and disproportionately higher increases among Asian, Black, and Latino/a communities (Dietz et al., 2020; Maenner et al., 2023). Several factors have driven these increases, including but not limited to **(a)** broader diagnostic criteria, **(b)** improved screening methods, **(c)** heightened public awareness, and **(d**) increased recognition of autistic individuals previously overlooked due to subtler presentations (Adreon & Durocher, 2007; Maenner et al., 2023).

However, despite these advances in identification, autistic adults face substantial barriers to meaningful employment, experiencing unemployment or underemployment at rates exceeding 58%, compared to approximately 5% for the general population (Nicholas et al., 2014; Ohl et al., 2017; Roux et al., 2023; Scott et al., 2018). When employed, many autistic individuals occupy positions that do not fully leverage their education, skills, or expertise, resulting in limited career advancement and inadequate workplace accommodations (Harvery et al., 2021; Nicholas et al., 2014; Roux et al., 2023; Taylor et al., 2015). The insufficiency of public support systems exacerbates this situation, with only 1.1% of autistic adults accessing critical services like Vocational Rehabilitation (VR) and Medicaid Home and Community-Based Services (HCBS), leaving approximately two million autistic adults without necessary employment support (Roux et al., 2023). These systemic gaps highlight a critical need to strengthen transition support structures facilitating autistic individuals' progression from higher education to meaningful employment.

Postsecondary education significantly improves employment opportunities and life satisfaction for autistic adults; however, only 44% of autistic individuals enroll in higher education, compared to 67% of the allistic population, presenting clear opportunities for improved access and institutional support to facilitate academic success (Anderson et al., 2017; Schembri-Mutch et al., 2023). Among autistic undergraduates, roughly 39% persist to graduation, underscoring the substantial potential for targeted interventions aimed at bolstering retention and completion rates (Petcu et al., 2021). Factors influencing these outcomes include self-determination, executive functioning, and, particularly, social connectedness (Petcu et al., 2021; Tomaszewski et al., 2022). Enhancing social engagement and fostering belonging positively impacts autistic students' well-being and persistence in academic settings, mitigating feelings of isolation and depression frequently reported by autistic undergraduates (Hamilton & Petty, 2023; Lei & Russell, 2021). Moreover, addressing institutional expectations that drive masking behaviors, where autistic individuals suppress neurodivergent traits to align with perceived social norms, can further empower autistic undergraduates by fostering self-determination and authentic engagement within higher education communities (Hamilton & Petty, 2023; Petcu et al., 2021).

Additionally, inclusive workforce development has become a critical priority, especially within science, technology, engineering, art, and mathematics (STEAM) fields, where autistic students enroll at nearly twice the rate of allistic peers but often exit prematurely due to inadequate institutional supports (Bakker et al., 2023; Crabtree et al., 2024). Interventions focusing on self-efficacy, collaborative skills, and tailored supports have successfully increased retention among historically marginalized groups, indicating promising strategies adaptable to autistic students in STEAM disciplines (Cobian et al., 2022; Daily & Eugene, 2013; Yassa & Chen, 2023).

Despite abundant literature identifying barriers to autistic students' successful participation in postsecondary education, fewer empirical studies have systematically evaluated interventions to

increase social engagement and campus connectedness. This research gap is particularly pressing given the documented relation between social support, academic persistence, and subsequent career opportunities (Duerksen et al., 2021; Raymaker et al., 2022). Although customized employment strategies and mentorship initiatives show promise in bridging higher education and equitable career outcomes, a distinct lack of integrated academic accommodations and socioemotional support interventions remains (Cheriyan et al., 2021; Duerksen et al., 2021). Therefore, it is imperative to empirically evaluate structured interventions, particularly using robust methodologies such as single-case experimental designs (SCED), that enhance self-determination, executive functioning, and social engagement skills among autistic students. Addressing this need promises improvements in degree completion rates and employment outcomes and meets broader societal imperatives to cultivate a diverse, inclusive workforce suited to a dynamic, global economy.

Systematic Literature Review

The growing prevalence of autism among postsecondary students underscores a critical need for evidence-based interventions that address their specific support needs. Although enrollment rates for autistic students in higher education have increased due to expanded screening and rising awareness, attrition rates remain markedly higher than those of allistic peers or peers with other disabilities. To investigate how best to support autistic students in postsecondary settings, I conducted a systematic review guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework. Specifically, this review examined SCED studies to (a) identify the types of interventions used to support autistic college students and their effectiveness, (b) explore perceived barriers to their participation and success, and (c) analyze the methodologies employed in the included studies.

In addressing these research questions, this review illuminates the systemic barriers autistic students encounter across multiple domains (e.g., social, academic, independent living) and the

interventions that show promise in mitigating these barriers. Findings indicate that targeted supports focusing on social skills training, technology-based self-management strategies, and structured organizational aids often yield positive outcomes in improving retention and overall well-being. However, gaps persist in terms of methodological rigor, long-term generalizability, and attention to underrepresented student populations and settings (e.g., community colleges, specific majors). Systematic reviews are essential in synthesizing the current state of research, guiding future inquiries, and informing institutional policies, thereby enhancing the capacity of postsecondary environments to support the academic and social success of autistic students.

Meta-Analysis

The systematic review provided comprehensive demographic data from existing studies, facilitating a meta-analysis that systematically evaluated outcomes and highlighted the importance of evidence-based interventions to support the increasing number of autistic students enrolling in postsecondary education. This approach underscores opportunities to improve persistence rates by addressing factors influencing attrition among autistic undergraduates. To systematically examine the effectiveness, I conducted a meta-analysis of SCED studies, following the PRISMA framework.

This meta-analysis aimed to: (a) identify the types of interventions implemented to support autistic college students, (b) evaluate their effectiveness through SCED studies, and (c) examine methodological characteristics and potential moderators influencing outcomes. I systematically extracted data from eligible SCED studies to construct a comprehensive dataset, incorporating variables such as funding source, visual analysis procedures, adherence to quality indicators, participant characteristics, and intervention components.

Findings revealed a moderate-to-large overall positive effect of interventions ($g = 0.83$), with significant heterogeneity present in the multivariate model. Further moderator analyses indicated that certain factors (e.g., methodological rigor, intervention structure) were correlated with more

substantial benefits. The analysis suggested potential publication bias, underscoring the need for continued rigorous research and transparent reporting to promote replication in future studies.

Single-Case Experimental Design

Findings from the systematic review and meta-analysis provided critical insights into social factors affecting autistic undergraduates in STEAM fields, informing the development of a SCED aimed at enhancing their academic success and sense of belonging during the transition to higher education. By systematically identifying key areas influencing persistence and attrition, this approach extends the existing literature and addresses opportunities for targeted intervention. In this study, I used a SCED to evaluate whether the Program for the Education and Enrichment of Relational Skills for Young Adults (PEERS®-YA) enhances social engagement and strengthens campus connectedness and overall positivity among autistic undergraduates. Specifically, I measured the constructs of belonging and positivity using a multiple baseline design combined with experience sampling methodology (ESM), systematically capturing real-time social interaction data. Additionally, I evaluated changes in participants' social-relational skills and perceptions of belonging using pre- and post-intervention assessments. By exploring the functional relation between enhanced social engagement and improved academic and social outcomes, I seek to inform the development of targeted supports for autistic college students.

Although preliminary visual analysis of daily ESM data suggested some therapeutic trends, particularly for one participant, these were not consistently verified as statistically significant using the conservative dual-criterion method for weekly aggregated data, leading to largely inconclusive results for these primary outcomes. However, pre-post assessment data indicated gains in overall campus connectedness for both participants, mixed results for social skills knowledge gains, and high social validity ratings, emphasizing the complexities of intervention response and the value participants place on the program structure.

Chapter 2: Literature Review

In 2017, the Centers for Disease Control (CDC) estimated the prevalence of autism among adults aged 18 to 84 years at 5.5 million individuals (Dietz et al., 2020). Furthermore, these estimations revealed a corresponding prevalence rate of approximately 1 in 6 (17%) among children aged 3 to 17 diagnosed with developmental disabilities between 2009 and 2017, with 1 in 54 children identified as autistic during this period (Dietz et al., 2020). Subsequent data from 2020 unveiled a noteworthy surge in autism diagnoses, with an estimated prevalence of approximately 1 in 36 (2.8%) among 8-year-old children (Maenner et al., 2023). Additionally, autism prevalence demonstrated marked increases among Asian, Black, and Hispanic children, rising by a minimum of 30% between 2018 and 2020. In contrast, prevalence among White children also exhibited growth, albeit at a rate of 14.6% during the same period (Maenner et al., 2023).

Contemporary research trends suggest increased rates of autism diagnoses can be attributed to a confluence of factors including, but not limited to, (a) the broadening of diagnostic criteria, (b) expanded screening initiatives, and (c) heightened public awareness (Adreon & Durocher, 2007; Maenner et al., 2023). These factors may contribute to the observed rise in diagnoses, particularly among underserved demographic groups, including autistic individuals whose autism may have been previously overlooked during their formative years due to subtler presentations or those who did not fit earlier, narrower diagnostic criteria (Adreon & Durocher, 2007; Maenner et al., 2023).

Numerous autistic individuals actively seek enrollment in postsecondary educational institutions and can achieve postsecondary degrees (Anderson et al., 2017; Schembri-Mutch et al., 2023). However, empirical evidence indicates that participation rates in postsecondary education among autistic individuals are significantly lower than those of the general population and their peers with speech or language impairments, learning disabilities, and/or cognitive disabilities. Specifically, enrollment rates for autistic students remain at 44%, in stark contrast to the 67% observed in the

general population and overall higher enrollment across peer disability categories, with the exception of individuals with intellectual and/or developmental disabilities or multiple disabilities classifications (See Table 1; Anderson et al., 2017; Schembri-Mutch et al., 2023). Nevertheless, contemporary advancements in screening methods, diagnostic practices, heightened awareness, and improved accessibility have collectively contributed to a notable increase in the enrollment of autistic students in postsecondary educational institutions, with current estimates suggesting that approximately 1.9% of students in higher education are autistic (Anderson et al., 2017; Dwyer et al., 2023). Despite this progress, it is essential to underscore that autistic students often have college preparation experiences in their secondary education that differ from traditional benchmarks, often opting for enrollment in 2-year colleges (32%) rather than 4-year institutions (17%).

Table 1. Postsecondary Enrollment By Disability Classification

	SLD	**SLI**	**IDD**	**E/BD**	**HI**	**VI**	**OI**	**OHI**	**ASD**	**TBI**	**MD**	**DB**
Enrollment:												
- any level	66.8	66.9	***28.7***	53.0	74.7	71.0	62.0	65.7	**43.9**	61.0	***32.8***	56.8
- 2-year	49.9	46.0	***18.9***	37.7	51.5	51.5	50.3	51.6	**32.2**	42.2	***21.7***	36.9
- trade	35.8	28.5	***16.4***	33.3	42.9	26.2	26.2	32.2	**21.0**	36.9	***17.5***	22.1
- 4-year	21.2	32.5	***6.7***	10.8	33.8	40.1	26.1	19.6	**17.4**	18.5	***7.4***	23.7

Note: *SLD = Specific Learning Disability; SLI = Speech-Language Impairment; IDD = Intellectual/Developmental Disabilities; E/BD = Emotional/Behavioral Disability; HI = Hearing Impaired; VI = Visually Impaired; OI = Orthopedic Impairment; OHI = Other Health Impairment; ASD = Autism Spectrum Disorder; TBI = Traumatic Brain Injury; MD = Multiple Disabilities; DB = Deaf-Blindness*

Furthermore, among the 44% of autistic students who do enroll in postsecondary education, completion rates remain strikingly lower (39%) in comparison to their peers in general education (59%) and their peers belonging to different disability categories (50%; Petcu et al., 2021). Although there exists a paucity of comprehensive research regarding the specific barriers and potential solutions to the higher attrition rates among autistic individuals in higher education, extant literature aligns with reports of compromised mental well-being attributed to several contextual factors. These factors include, but are not limited to, (a) the presence of the "hidden curriculum" inherent to university life, which encompasses norms related to behavior, study habits, and social interactions; (b) a negative perception of education stemming from unsupportive primary- and secondary-school

environments; (c) challenges in establishing a sense of belonging, often arising from overwhelming and inescapable social situations with peers, which may engender incidents of bullying or feelings of isolation; and (d) the phenomenon of masking, wherein autistic individuals suppress their innate neurodivergent responses to mitigate negativity, leading to feelings of exhaustion, dissociation, psychological distress, and a heightened risk of dropping out before attaining their degrees (Hamilton & Petty, 2023; Petcu et al., 2021).

Prior Research

Postsecondary Supports for Allistic Peers

Traditional Students. Traditionally enrolled postsecondary students frequently experience significant stressors impacting academic performance, persistence, and satisfaction. Factors contributing to elevated stress levels include academic demands, financial pressures, interpersonal relationships, and career uncertainty, which can negatively affect students' psychological well-being, physical health, and academic outcomes (Lisnyj et al., 2023). Effective institutional supports identified to mitigate these stressors involve strengthening intrapersonal resources through interventions targeting self-efficacy, resilience, resolve, academic motivation, and psychological well-being. Additionally, promoting interpersonal resources through enhanced social capital (e.g., fostering strong relationships with peers and faculty, structured mentorship programs, mindfulness training, robust student support systems) has proven beneficial in reducing perceived stress and improving academic success. Emphasizing comprehensive and proactive support mechanisms can significantly enhance traditionally enrolled students' resilience to stress, contributing positively to their overall educational experiences and success (Lisnyj et al., 2023).

Nontraditional Students. Nontraditional students, characterized by factors such as older age, part-time enrollment, full-time employment, caregiving responsibilities, military experience, delayed college entry, and first-generation college status, constitute a growing majority in

postsecondary education (MacDonald, 2018). They encounter unique barriers, including difficulties managing work-life-school balance, anxiety, addressing varied levels of academic preparedness due to diverse prior educational pathways, and challenges adapting to contemporary educational technologies and methodologies. Effective interventions supporting nontraditional students emphasize early and mandatory remedial and seminar courses, flexible scheduling including evening, weekend, and online course options, differentiated instructional approaches integrating practical and experiential learning, and strong, supportive instructor-student relationships emphasizing personal connection, structured feedback, and positive reinforcement (MacDonald, 2018).

Low-Income, High-Potential Students. Low-income students experience notable disparities in postsecondary educational outcomes, facing significant barriers in college enrollment, persistence, and degree attainment compared to their more affluent peers (Renbarger & Long, 2019). Systematic research identifies several promising interventions aimed at supporting this student population, including summer bridge programs that enhance college readiness, advanced coursework such as Advanced Placement (AP) or International Baccalaureate (IB) programs, dual-credit opportunities, comprehensive financial aid support, targeted dissemination of college-related information, and ongoing mentorship or year-round support programs. Notably, these interventions not only increase access to college but also positively impact students' persistence, academic success, and non-cognitive outcomes such as peer support, confidence, and social integration. Despite these benefits, implementation inconsistencies and limited support resources continue to constrain their effectiveness, underscoring the necessity of strategically combining multiple support elements to maximize outcomes for low-income students (Renbarger & Long, 2019).

Postsecondary Students with Disabilities

Recent systematic reviews reveal consistent themes in promoting postsecondary success for students with disabilities, particularly in transition planning, self-advocacy, and robust institutional

supports (Madaus et al., 2024; Nail, 2024). Nail (2024) underscores student-centered factors, including but not limited to (a) knowledge of one's disability, (b) awareness of available accommodations, and (c) the development of self-determination, as foundational to student success, explicitly connecting these factors to crucial social interactions with peers, faculty, and institutional personnel. Madaus et al. (2024) echo these priorities, emphasizing the centrality of relational supports and social experiences while advocating for more rigorous empirical research to validate existing supports across diverse institutional contexts. Madaus et al. (2024) and Nail (2024) highlight the importance of social domains in fostering students' self-advocacy and relational skills, which are integral to successfully navigating postsecondary environments. These studies simultaneously indicate increased attention to the multifaceted experiences of students with disabilities participating in higher education while underscoring the need for evidence-based interventions to enhance transitional readiness, self-advocacy, and institutional responsiveness within supportive social contexts to better serve diverse student populations.

Postsecondary Students with Autism

Nelson et al. (2023) conducted a scoping review that highlights critical barriers and supports influencing the postsecondary experiences of autistic students. Despite increasing enrollments, autistic students face substantial challenges in receiving autism-specific supports to address perceived barriers, including but not limited to (a) significant mental health concerns, (b) challenges in navigating social interactions and communication differences within typical campus environments, and (c) environmental stressors related to sensory sensitivities (Nelson et al., 2023). Supports identified as effective contributors to success include tailored academic accommodations, structured peer mentorship programs, specialized transition supports such as pre-entry orientation, and readily available mental health services. Conversely, substantial barriers encompass inadequate individualized accommodations, lack of awareness or misinformation about autism among faculty

and peers, restrictive documentation requirements for disability-related supports, and challenges navigating institutional support services. Additionally, disclosure emerges as a complex issue; while necessary to access accommodations, it simultaneously exposes autistic students to stigma and social isolation risks, underlining the urgent need for broader campus-wide initiatives promoting understanding, inclusivity, and flexible individualized support structures (Nelson et al., 2023).

Current Review

The primary objective of my systematic review was to examine interventions designed to support autistic students in postsecondary settings, specifically focusing on studies that implemented a single-case experimental design (SCED). My secondary goals included, but were not limited to, (a) identifying frequently cited barriers to postsecondary participation for autistic students, (b) assessing interventions in relation to these perceived barriers, and (c) reviewing the varied methodologies implemented in the research. To support these objectives, I posed the following three research questions:

1. What types of interventions have been implemented to support autistic students in postsecondary settings using a single-case experimental design?
2. What are the most reported perceived barriers that autistic students face in postsecondary education, and how do existing interventions address these barriers?
3. How do the single-case experimental design methodologies employed in postsecondary intervention studies for autistic students vary, and what are the implications of these methodologies on the reliability and applicability of the results?

Method

Review Methods

This review was informed and guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for conducting systematic reviews, specifically implementing the updated guidelines and exemplars (Page et al., 2021).

Search Strategy

I adopted a systematic approach in formulating the search strategy, comprising the following steps: (a) scrutinizing extant literature on the subject, (b) discerning keywords from the reviewed articles, (c) constructing a preliminary Boolean string derived from identified keywords, (d) executing an initial search deploying the Boolean string and its variations within the title and abstract to establish comprehensive phrasing, and (e) systematically examining the titles and abstracts of resulting articles to ascertain their relevance to the search terms. To encompass all potential variations, acronyms (e.g., ASD, ADHD, IPSE, CTP) and colloquial terminology (e.g., Asperger's) were deliberately incorporated into the Boolean string.

My search strategy involved querying two renowned databases, namely the *Education Resources Information Center* (ERIC) and *ProQuest Dissertations and Theses Global*. Initially, I used the Boolean string to search the abstract of documents: ("neurodiversity" OR "neurodiverse" OR "autism spectrum" OR "autism spectrum disorder" OR "ASD" OR "autistic" OR "Asperger's" OR "Asperger's syndrome" OR "attention deficit disorder" OR "ADD" OR "attention deficit hyperactivity disorder" OR "ADHD" OR "dyspraxia" OR "dyslexia" OR "dyscalculia") AND ("postsecondary education" OR "post-secondary education" OR "higher education" OR "college" OR "university" OR "vocational rehabilitation" OR "career training" OR "technical training" OR "career and technical training" OR "CTE" OR "occupational training" OR "short-term training"

OR "inclusive education" OR "inclusive postsecondary education" OR "IPSE" OR "comprehensive transition and postsecondary program" OR "CTP").

To optimize my search strategy, I limited results to dissertations and peer-reviewed articles published from 1990 onward, coinciding with the "second wave" of federal transition policy driven by the Individuals With Disabilities Act of 1990 (PL 101–476), which for the first time mandated (a) inclusion of transition services in individualized education programs, (b) formal linkages with outside agencies, and (c) state- and local-level monitoring and reporting of student outcomes (Morningstar et al., 2012). Morningstar and colleagues (2012) note that this policy shift, including the related momentum from the School-to-Work Opportunities Act of 1994, sparked a surge of empirical investigations into perceived barriers and interventions throughout the 1990s, making 1990 the most defensible threshold for capturing the whole arc of transition-services research. ERIC-specific documents served as an exception to this criterion. The revised search, conducted on May 16, 2023, targeted titles and abstracts using the following Boolean string: ("neurodiversity" OR "neurodiverse" OR "autism spectrum" OR "autism spectrum disorder" OR "ASD" OR "autistic" OR "Asperger's" OR "Asperger's syndrome" OR "attention deficit disorder" OR "ADD" OR "attention deficit hyperactivity disorder" OR "ADHD" OR "dyspraxia" OR "dyslexia" OR "dyscalculia") AND ("postsecondary education" OR "post-secondary education" OR "higher education" OR "college" OR "university" OR "vocational rehabilitation" OR "career training" OR "technical training" OR "career and technical training" OR "CTE" OR "occupational training" OR "short-term training" OR "inclusive education" OR "inclusive postsecondary education" OR "IPSE" OR "comprehensive transition and postsecondary program" OR "CTP").

Ancestral Search

An ancestral search was conducted for all documents that met inclusion criteria through the electronic search process (n = 12). Each reference item was hand searched, with titles that met the

inclusion criteria receiving further screening in the abstract section. This process resulted in three additional articles being included in the final count, bringing the total to 15 articles.

Inclusion/Exclusion Criteria

The following inclusion criteria were applied to identify documents, with iterations of the screening process informed by slight alterations to my dissertation focus (See Figure 1). In the initial screening, my focus included four primary inclusion criteria:

- **(a)** studies were published in English, considering my formal training and expertise in research methods and terminology, which are limited to my native language, and I lacked sufficient resources to ensure accurate translation of non-English articles;
- **(b)** studies that qualified as a research report, defined as peer-reviewed publications and completed dissertations presenting original empirical findings based on direct data-collection via quantitative and/or qualitative methodologies;
- **(c)** studies that included neurodivergent participants, defined as "a concept that regards differences in behavior and brain function as part of the normal variation of the human population," including conditions such as autism spectrum disorder, attention-deficit/hyperactivity disorder, and other neurodiverse classifications (e.g., dyslexia, dyspraxia, dyscalculia, dysphonia; Fung, 2021). During this stage of screening, I included any articles who explicitly stated neurodiversity or neurodiverse conditions in the title or abstract;
- **(d)** studies that were conducted in postsecondary contexts, defined as "the range of formal learning opportunities beyond high school, including those aimed at learning an occupation or earning an academic credential" (Institute of Education Sciences, n.d.).

In the second screening, I restricted the pool of articles by requiring:

(e) studies which investigated participants with autism spectrum disorder, including articles that explicitly stated the population in the title or abstract. At this stage, I decided to focus on studies that investigated perceived barriers to postsecondary participation, excluding articles that allowed self-reported diagnoses since these students were not likely to register with disability services.

During the third screening, I opted to focus on traditional enrollment, including:

(f) studies that included participants enrolled at a four-year institution, excluding articles that included individuals living in off-campus settings.

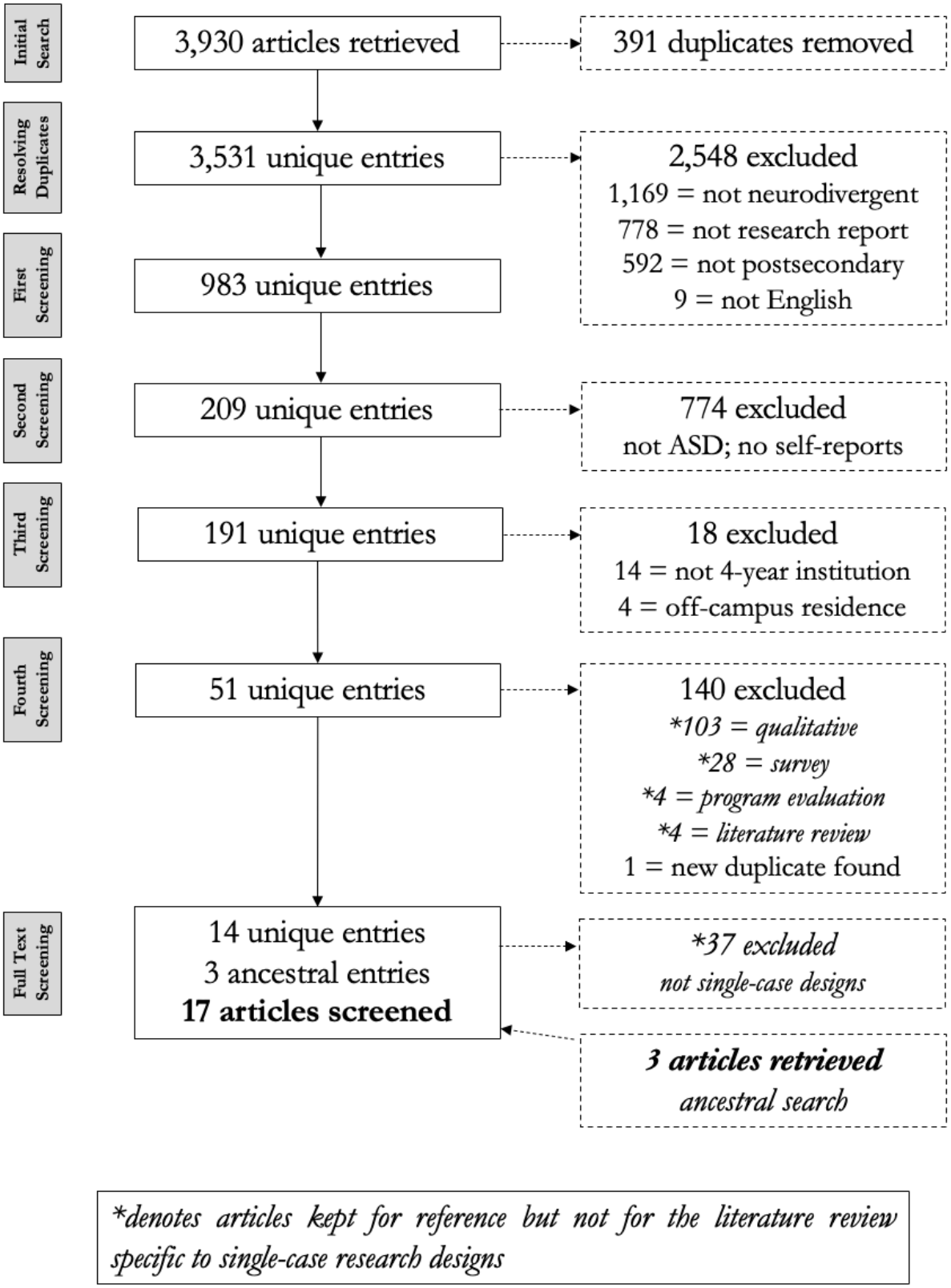
For the final screening, I decided to implement quantitative research for my dissertation, specifically single-case experimental designs, and excluding:

(g) studies that implemented single-case experimental designs that did not meet minimum NTACT Quality Standards (i.e., AB design), studies that relied solely on qualitative methodologies, and studies that performed surveys, program evaluations, or literature reviews.

Screening Procedure

To determine appropriate studies for data extraction, I compiled and exported database searches into batch files (i.e., specifically RIS files), which were then imported into *Rayyan* (rayyan.ai), a free-to-use online service for organizing and managing collaborative systematic literature reviews. After uploading the files to the platform, I screened the titles and abstracts of the initial pull of 3,531 articles, after removing duplicates, according to decision parameters detailed in the PRISMA flow chart (See Figure 1).

Figure 1. Screening Diagram



Interobserver Agreement on Initial Screening

A second independent reviewer participated in screening to document agreement in the screening process. This individual is a doctoral-level scholar with a Master's in Psychological Sciences, whose research interests include the influence of social identities on perception and how such perceptions impact student motivation, engagement, and academic emotions. Prior to the interobserver agreement (IOA) screening, the second reviewer underwent training to ensure alignment with the study's screening procedures. This training included familiarization with the *Rayyan* (rayyan.ai) systematic review platform, a thorough review and clarification of the inclusion and exclusion criteria, and an overview of the labeling protocols to ensure consistent application of the eligibility criteria.

For the IOA evaluation, 20% of the initial article pool (See Table 2; $n = 706$) was randomly assigned to the second reviewer via the Rayyan platform's randomization feature. The second reviewer independently applied the inclusion and exclusion criteria to these titles and abstracts, with resulting decisions compared to those of the primary reviewer. Of these double-screened articles, 44 were initially coded differently between the two reviewers, yielding an IOA >90%. Conflicting screening decisions were subsequently resolved through discussion. These discussions focused on refining the screening process and led to clarifications in four key domains: (a) the inclusion and exclusion criteria were made more explicit to ensure replicability; (b) duplicate records that had been inadvertently retained were identified and removed; (c) nuanced interpretations of certain abstracts were deliberated to reach consensus; and (d) a few mislabeled articles were identified and corrected.

For full-text screening, I independently reviewed all articles ($n = 51$) due to time constraints and scheduling conflicts, so IOA could not be calculated at this stage. To mitigate potential bias, I followed the predefined inclusion criteria, SCEDs with potential to meet the minimum NTACT Quality Indicators and kept detailed records of every exclusion decision (See Figure 1). Once the

pool was finalized (n = 15), I provided training for a departmental colleague, a doctoral scholar with a Master's in School Counseling and research interests in postsecondary supports for neurodiverse students, which included a thorough review of the predefined inclusion criteria of SCEDs. I then asked this colleague to screen 33.3% (See Table 2; n = 5) of these articles. Our inclusion judgments achieved 100% IOA, confirming consistency in the final selection process. I acknowledge the absence of IOA during the initial full-text review as a methodological limitation and will discuss it further in the section on **Limitations**.

Table 2. Interobserver Agreement (Article Pool)

Screening Phase	Total	Agreement	Disagreement	Percentage
- Initial Pool	706	662	44	93.8%
- Final Pool	5	5	0	100%

Quality Assessment

To systematically evaluate the internal validity of the included full-text articles, I used the *National Technical Assistance Center on Transition: The Collaborative* (NTACT:C) *Quality Indicator Checklist for Single-Case Experimental Designs*. NTACT:C is a technical assistance center co-funded by the United States Department of Education's Office of Special Education Programs and the Rehabilitation Services Administration. It provides information, tools, and supports to assist stakeholders in delivering effective services and instruction for secondary and postsecondary youth with disabilities. Given the context of this systematic review and NTACT:C's dedication to transition supports and services, these quality indicators were the most appropriate choice for assessing study quality. I independently evaluated the design quality of each included full-text article using this checklist.

To ensure reliability in the quality assessment process, I trained a second independent screener, a doctoral scholar with a Master's in School Counseling and research interests in postsecondary supports for neurodiverse populations. During training, we jointly reviewed the NTACT:C Quality Indicators and completed a practice assessment on a separate article to calibrate

our application of the criteria. All full-text articles were uploaded to a shared *OneDrive* folder to facilitate collaboration. The second screener then randomly selected approximately 33.3% (n = 5) of the included studies from this folder for independent quality assessment. Using the same checklist, the second screener evaluated the design quality of these studies. I then calculated IOA based on this subset of double-coded studies to quantify the consistency between our independent assessments.

NTACT:C Quality Indicators

The single-case quality indicators were developed through a collaborative effort among subject-matter experts in special education, rehabilitation services, vocational training, and research methodology. Drawing on foundational standards articulated by Horner et al. (2005) and Kratochwill et al. (2013), this team translated evidence-based practice criteria into a discrete, observable checklist that maps directly onto key aspects of internal validity, including but not limited to the precision of participant and setting descriptions, systematic manipulation of the intervention, and repeated measurement with reliability data. Subject-matter experts then refined the language and minimum thresholds (e.g., at least three demonstrations of effect, ≥ 80% IOA) to ensure that the indicators balanced methodological rigor with practical relevance for transition-focused research. This consensus-driven process resulted in a transparent framework for classifying single-case studies as "high," "acceptable," or "did not meet" quality standards, based on their adherence to these standardized criteria.

Initial Criteria. Before determining overall quality, all single-case experimental designs must meet the minimum initial criteria to be considered for further coding. Initial criteria include (a) dependent variable(s) focused on a transition-related skill (e.g., life skills, functional academics, employment skill) and (b) participants include youth and young adults with disabilities aged 18 to 26. Furthermore, rigorous designs include (a) reversal, (b) multiple baseline, (c) multiple probe, (d)

changing criterion (if including reversal), and (e) alternating treatment (if including baseline phase). AB designs do not meet the quality standards outlined by NTACT.

Participant. (1) Participants are operationally described with sufficient detail, allowing others to select individuals with similar characteristics for potential replication. (2) Processes for selecting participants are operationally described with replicable precision, allowing external researchers to duplicate procedures.

Setting. (3) Critical features of the physical setting are detailed with sufficient precision, allowing replication, with features of the setting operationally defined in a form that external researchers could recruit similar participants who inhabit similar settings.

Dependent Variable/Measure. (4) Dependent variables are operationally defined, with each dependent variable described for valid, consistent assessment of the variable. (5) Each dependent variable is measured with a procedure that generates a quantifiable index (e.g., frequency, count). (6) Measurement processes are described with replicable precision. (7) Dependent variables are measured repeatedly across time, allowing for observation of patterns across conditions or phases. (8) Data is collected on the reliability or IOA associated with each dependent variable, with IOA levels meeting minimum standards (e.g., IOA = 80%; Kappa = 60%).

Independent Variable/Intervention. (9) Independent variable is operationally defined to allow both valid interpretation of the results and accurate replication of the procedures, including but not limited to descriptions of materials and specific actions, while avoiding vague, generic descriptions susceptible to high variability in implementation. (10) The Independent variable is systematically manipulated under the researcher's control, who determines when and how the intervention will be implemented. (11) Overt measurement of the fidelity of implementation for independent variables, with documentation of procedural fidelity measures.

Procedures. (12) A baseline phase that provides repeated measurement of a dependent variable, establishing a pattern of responding that is used to predict the pattern of future performance without introducing the independent variable, with five or more responses recommended and fewer permitted if the pattern was established. (13) Procedural characteristics of the baseline conditions are described with replicable precision, with an equivalent level of detail as the treatment (or intervention) phase to allow for comparisons and replication.

Design/Graph/Results. (14) The design provides a minimum of three demonstrations of experimental effect at different points in time, either across different points in time for one participant or across a minimum of three participants. A demonstration of effect is an increase, decrease, or desired reversal in direction of the anticipated pattern of baseline data. While graphs can provide evidence, a functional relation is compromised when (a) there is significant latency between manipulation of the independent variable and the dependent variable; (b) mean changes across conditions are similar to changes within conditions; or (c) trends do not follow those predicted when introducing the intervention. (15) The design controls for common threats to internal validity, demonstrated through (a) introduction and withdrawal of the independent variable, (b) staggered introduction of the independent variable, or (c) manipulation of levels of the independent variable across observation periods. (16) Experimental effects are replicated across participants, settings, or materials to establish external validity, enhanced through replicable descriptions of (a) participants, (b) study context, and (c) factors influencing behavior prior to intervention. This is further enhanced through the use of multiple participants or settings and multiple measures of the dependent variable, demonstrated through systematic replications across multiple locations and researchers.

Social Validity. (17) The dependent variable is socially important, with the included participant(s) completing a measure to determine if the dependent variable is perceived as important and detailed in the manuscript. (18). The magnitude of change in the dependent variables resulting

from the intervention is measured as socially important, with the amount of increase or decrease in the targeted behavior as a result of independent variable manipulation being considered meaningful. (19) The author describes the implementation of the independent variable as practical and cost-effective, providing detailed descriptions of how the benefits of the intervention outweigh the potential costs (e.g., time, cost, personnel). (20) Social validity is enhanced by the implementation of the independent variable over extended periods, enhanced by demonstrations in contexts that are not overly disruptive to regular routines.

Quality Determination. Studies that satisfied all 20 indicators were designated as "High Quality." Those that met indicators 1-16 and at least one of indicators 17-20 were designated as "Acceptable Quality." Any study failing to meet one or more of the core indicators (1-16) and none of the social validity or practicality indicators (17-20) was classified as "Did Not Meet Quality," with the specific indicator numbers recorded for transparency. This approach ensured a consistent, rule-based determination of study rigor across the review.

Interobserver Agreement of Quality Appraisal

I evaluated all included studies (n = 15) using the NTACT:C Quality Indicators, summarizing adherence according to each indicator domain (See Appendix A: Table 4). Regarding participant criteria (Indicators 1-2), all studies (100%) provided adequately detailed descriptions of participants and operationalized their participant selection procedures. For the setting description (Indicator 3), 14 of the 15 studies (93%) thoroughly detailed critical setting features, while one provided only partial information.

In terms of measurement rigor (Indicators 4-8), the results indicated consistent adherence. Fourteen studies (93%) operationalized their dependent variables (Indicator 4). In comparison, all 15 studies (100%) employed quantifiable measurement procedures (Indicator 5), offered replicable

measurement protocols (Indicator 6), repeatedly measured variables across conditions (Indicator 7), and reported reliability or IOA at least once (Indicator 8).

Regarding the definition and fidelity of the intervention (Indicators 9-11), adherence was strong, with all 15 studies (100%) providing precise and replicable descriptions of the intervention (Indicator 9) and systematically controlling for the implementation of the intervention (Indicator 10). However, procedural fidelity checks (Indicator 11) were less consistently documented, with only 10 studies (67%) explicitly reporting fidelity measures.

In assessing procedures and design rigor (Indicators 12-16), 14 studies (93%) included rigorous baseline phases with adequate data points and detailed procedural descriptions (Indicators 12–13). Similarly, 14 studies (93%) demonstrated at least three apparent experimental effects (Indicator 14), controlled alternative explanations (Indicator 15), and replicated intervention effects across participants, settings, or materials (Indicator 16).

Adherence to the social validity and sustainability indicators (17-20) varied more substantially. Twelve studies (80%) addressed the social importance of targeted behaviors (Indicator 17), and 14 studies (93%) evaluated the social significance of behavior change (Indicator 18). However, the practicality and cost-effectiveness of interventions (Indicator 19; 33%) and the extended, naturalistic implementation by typical intervention agents (Indicator 20; 60%) were less frequently documented.

After conducting my initial quality evaluation, I involved the same departmental colleague who had previously confirmed the SCEDs to independently code a random subset (33.3%; $n = 5$) of the studies for IOA purposes (See Table 3; See Appendix A: Figure 1 for Full IOA). Prior to independent coding, we completed joint training, which included a thorough review of the NTACT:C Quality Indicators and collaborative coding of a separate article not included in the final pool. IOA at the study level ranged from 85% to 100% (M = 91%, SD = 6.52%), whereas IOA at

the individual indicator level ranged from 40% to 100% (M = 91%, SD = 16.51%). The most considerable discrepancies were identified in Indicators 19 (practicality and cost-effectiveness) and 20 (extended, naturalistic implementation). Upon reviewing these discrepancies, I observed that authors often documented social validity evaluations separately from descriptions of practicality and naturalistic implementation, contributing to interpretive ambiguity. To address these disagreements, we reconvened, revisited the articles, and resolved each discrepancy through consensus discussions.

Table 3. Interobserver Agreement (Quality Assessment By Indicator)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.8	0.8	1	0.8	1	1	1	1	1	1	1	1	1	1	1	1	1	0.8	0.6	0.4

Study Characteristic Coding

To systematically summarize the characteristics of the studies included in this literature review, I developed and applied a structured coding guide informed by the *NTACT:C Quality Indicator Checklist*, as outlined in the Quality Assessment section. Specifically, I based my coding categories on the checklist's domain framework, which encompasses key study elements, including participants, setting, dependent and independent variables, procedures, findings, and social validity indices. I applied the definitions and descriptions of each domain to ensure consistent data extraction across all studies. Notably, because I conducted the study characteristic coding independently, it was imperative to follow a clearly defined coding scheme to maintain consistency and reliability. This structured approach provides a transparent bridge between the screening narrative and the results overview, directly informing the organized presentation of participant demographics, study designs, and key variables in the next section.

Results

Following the initial search on May 16, 2023, a total of 3,930 articles were identified, which was subsequently refined to 3,531 after the removal of duplicate entries (n = 391). During the inaugural screening phase, researchers systematically excluded articles failing to adhere to specific

inclusion criteria: (a) publication in the English language, (b) classification as a research report, (c) occurrence within a postsecondary educational setting, and (d) inclusion of neurodivergent participants. This initial screening process yielded 983 articles deemed suitable for in-depth examination, while 2,548 articles were deemed ineligible for further consideration, failing to meet the specified minimum criteria. In the subsequent screening phase, the 2,548 excluded articles were subjected to further scrutiny, with reasons for exclusion outlined as follows: (a) absence of neurodivergent populations (n = 1,169), (b) not identified as a research report (n = 778), (c) lack of postsecondary settings (n = 592), or (d) not published in English (n = 9).

In the second round of screening, I chose to refine the focus to autistic college students with a formal diagnosis, eliminating the option for self-reported classification. This screening phase yielded 209 unique entries, excluding articles that did not involve participants with a confirmed autism spectrum disorder diagnosis (n = 774). For the third round of screening, researchers opted to further narrow the focus by excluding studies that featured participants in a community college setting (n = 14) or those who did not reside on campus (n = 4), resulting in 191 unique entries.

At this juncture, I opted to exclude articles that (a) centered exclusively on qualitative research (n = 103), (b) reported survey data from cross-sectional studies (n = 28), (c) constituted program evaluations (n = 4), or (d) was as systematic literature reviews (n = 4) during the fourth round of screening. Although these articles were excluded from consideration, they were retained for potential reference material. Additionally, during this stage, I identified a previously unnoticed duplicate that had not been identified in earlier phases on account of differences in author order. This reduction resulted in 51 unique documents, prompting the final stage of title and abstract screening by excluding any studies that did not use a single-case experimental design (n = 37).

Researchers conducted full-text screening on the final 14 articles, leading to the exclusion of additional articles (n = 2). One of the published articles originated from the author's dissertation,

which was consequently excluded from consideration but preserved for posterity (Ashbaugh et al., 2017). Another article employed an AB design without any means of replication within the study and failed to meet the minimum *NTACT:C Quality Indicators* for consideration (Gunn et al., 2017). During the full-text review, I observed that some documents included community college students, prompting a revisit to the initial exclusion of community college settings (n = 14). Upon closer examination, one of the previously excluded studies, which was conducted in a community college setting, met the additional inclusion criterion of adhering to the standards of a SCED, resulting in a final count of 12 unique entries for inclusion in the literature review, with the ancestral search adding three entries for a final total of 15 included articles and dissertations.

Study Characteristics

Participant Demographics

Across the analyzed studies (n = 15), a total of 45 participants were included (See Appendix A: Table 2). The majority of studies reported participant ages within the transition range (i.e., 18 to 26), with one study (Mann & Karsten, 2020) not reporting participant ages and another study (Koegel et al., 2016a) including participants aged 26 and above. Regarding gender identity, seven of the studies included only male or female participants, while eight studies included both genders among participants, with 29% (n = 13) identifying as female and 71% (n = 32) identifying as male. Among the hybridized studies, six demonstrated a preponderance of male participants, with two studies including more female subjects (Ashbaugh et al., 2017; McMahon et al., 2016).

When considering ethnic identity, seven studies (47%) did not report this demographic information. Of the remaining eight articles, every research study included at least one participant who identified as White (i.e., terms used by researchers included Caucasian, European, or European-American). Among the other noted ethnic identities, two studies included individual participants (4.4%) identifying as Asian-American participants (Engstrom, 2019; Koegel et al., 2016a), two

studies included single participants (4.4%) identifying as Latino/a (Koegel et al., 2013; Mason et al., 2012), and one study included one student (2.2%) identifying as Middle Eastern (Koegel et al., 2016b). Every study included autistic individuals, with two studies (4.4%) including definitions for Asperger's syndrome (Mann & Karsten, 2020; Mason et al., 2012). Additional disability categories included a comorbidity with attention-deficit/hyperactivity disorder (1 [2.2%]; ADHD; Engel, 2009) and separately identified intellectual disabilities (2 [4.4%]; McMahon et al., 2016; Wright et al., 2022).

Research Designs

Twelve of the included studies (80%) implemented some form of multiple-baseline design. Ten of the studies (67%) used a multiple-baseline design across participants, and two of the designs (13%) applied a multiple-baseline design across behaviors. Two of the studies (13%) performed multiple-probe designs, with one looking across behaviors (McMahon et al., 2016) and the other across participants (Wright et al., 2022), who further embedded a reversal design into the intervention. In one study (Gunn et al., 2017), researchers introduced an AB intervention across settings (7%); however, there is no discernible replication within the study.

Settings

All researchers collected data and introduced interventions in the natural setting, with every study occurring on a traditional university campus, university campus supported with an inclusive postsecondary education (IPSE) program, or a community college. Three of the studies included participants from multiple iterations of the aforementioned settings (Ashbaugh et al., 2017; Detar, 2020; Engstrom, 2019).

Dependent Variables

Seven of the studies included measures investigating natural occurrence of a social or social-adjacent behavior or skill. Five assessed social engagement skills, specifically monitoring social inquiry (Detar, 2020), positive-negative statements, affect, and interest (Koegel et al., 2016a),

empathic responses (Koegel et al., 2016b), correct-incorrect responses (Mann & Karsten, 2020), and reciprocal conversation (Mason et al., 2012). Two similar studies looked at social activities and peer engagement on campus (Ashbaugh et al., 2017; Koegel et al., 2013). Four articles assessed academic skills across multiple domains, including essay writing (Jackson et al., 2018), scientific vocabulary development (McMahon et al., 2016), note-taking (Reed et al., 2016), and on-task behavior and course engagement (Siko, 2018). Two studies evaluated independent living skills, with one monitoring punctuality and completion of appointments (Wright et al., 2022) and another assessing the percentage of completed daily living skills (Engstrom, 2019). The remaining two studies had relatively unique dependent measures in their designs, with Engel (2009) assessing the effects of repetitive behaviors, attention, and duration during therapy sessions, and Gunn et al. (2017) evaluating classroom management behaviors in a prospective teacher.

Independent Variables

Among the seven studies targeting social or conversational engagement and planning, four implemented a form of either video modeling or feedback connected to developing social or conversational skills (Detar, 2020; Koegel et al., 2016a; Koegel et al., 2016b; Mason et al.,2012). Detar (2020) focused on social inquiry training, specifically through conversational practice with neurotypical peers. Koegel et al. (2016a) included roleplay scenarios concurrently with video recording and feedback sessions to reframe negative statements, including provisions for participants to self-monitor their behaviors. In their additional study, Koegel and colleagues (2016b) included a visual framework during intervention procedures. The interventions from Mason et al. (2012) included verbal cues during the sessions. Two of the social or conversational studies (Ashbaugh et al., 2017; Koegel et al., 2013) looked primarily at structured social planning, including interventions for (a) motivational interests, (b) social activity options, (c) organizational skills, (d) peer mentors, and (e) social skills training. Koegel and colleagues (2013) included weekly meetings

and a systematic fading of supports for participants. In the study conducted by Mann and Karsten (2020), the intervention focused on honing conversational competency across multiple domains, including (a) problem detection, (b) solution provision, and (c) self-questioning training, with peer-led probes throughout each phase and natural environment tests.

Of the remaining studies, three included technology in their designs. Two studies (Siko, 2018; Wright et al., 2022) implemented wearable devices in their interventions. The study conducted by Siko (2018) included software that fed prompts to participants to self-monitor their targeted behaviors. Wright et al. (2022) focused on programmed reminders for participants to complete their designated daily tasks. McMahon et al. (2016) used *Aurasma*, an augmented reality application, to improve student understanding of scientific vocabulary, specifically allowing students to identify objects via their mobile device's camera. The remaining five studies had scant similarities in their interventions, including (a) animal-assisted therapy (Engel, 2009), (b) daily living skills checklist in tandem with peer mediation (Engstrom, 2019), (c) behavior skills training (i.e., social pragmatics, executive functioning; Gunn et al., 2017), (d) writing interventions (i.e., DATE, SRSD; Jackson et al., 2018), and (e) split-page note-taking (Reed et al., 2016).

Findings

Success within studies was relatively consistent, especially with interventions targeting social engagement. Within those specific studies, reports indicated mostly therapeutic trends across targeted measures (Ashbaugh et al., 2017; Detar, 2020; Koegel, 2013; Koegel, 2016a; Koegel, 2016b; Mann & Karsten, 2020), with some researchers measuring collateral data that showed additional improvements in (a) academic performance, (b) overall satisfaction with campus experiences, and (c) comorbid conditions (e.g., depression, anxiety; Ashbaugh et al., 2017; Detar, 2020; Koegel, 2013; Koegel, 2016a; Koegel, 2016b).

Within the other studies, researchers observed an improvement in their targeted behaviors or skills (Engstrom, 2019; Jackson et al., 2018; McMahon et al., 2016; Reed et al., 2016; Siko, 2018; Wright et al., 2022). Few studies either demonstrated no observed effect (Engel, 2009) or varied results within participants (Gunn et al., 2017; Mason et al., 2012).

Social Validity

In evaluating social validity specifically, most of the included studies incorporated measures to assess the acceptability and perceived value of the interventions. Four studies (27%; i.e., Ashbaugh et al., 2017; Engstrom, 2019; Jackson et al., 2018; Koegel et al., 2016b) utilized pre- and post-intervention questionnaires, revealing meaningful gains beyond quantitative behavioral improvements, such as increased satisfaction with peer interactions and overall campus experiences. Other researchers (n = 6 [40%]) relied primarily on post-intervention surveys or interviews (e.g., Detar, 2020; Engel, 2009; Koegel et al., 2013; Koegel et al., 2016a; McMahon et al., 2016; Wright et al., 2022), often reporting collateral benefits that included increased participant confidence, social engagement, and overall quality of life. Notably, two studies (Mann & Karsten, 2020; Siko, 2018; 13%) included evaluations by peers naïve to the research, adding an additional layer of objectivity to their social validity assessments. However, three studies (Gunn et al., 2017; Mason et al., 2012; Reed et al., 2016; 20%) did not include any measures of social validity, representing a limitation regarding the perceived practical significance and acceptability of their reported interventions.

Discussion

Autistic individuals are enrolling at institutions of higher education at increased rates, underscoring the critical need for support programs that facilitate their successful transition and retention in postsecondary settings (Nelson et al., 2023). The findings from this review suggest that while interventions, particularly those targeting social engagement and conversational skills, yield promising improvements in campus experience, academic performance, and overall quality of life,

significant gaps remain in the literature. Notably, preventative intervention studies that might reduce the high attrition rates observed among autistic undergraduates and graduates are scarce. This dearth of research deprives autistic students of the comprehensive support necessary for their academic and social success, especially when compared to the proactive and multifaceted supports typically offered to neurotypical peers. Effective institutional supports for traditionally enrolled students frequently include structured mentorship programs, interventions targeting resilience and self-efficacy, and robust social and academic support systems (Lisnyj et al., 2023). Nontraditional students similarly benefit from flexible scheduling, differentiated instructional methods, and strong instructor-student relationships emphasizing personal connection and structured feedback (MacDonald, 2018). Similarly, low-income, high-potential students often receive support through summer bridge programs, mentorship, dual-credit opportunities, and comprehensive financial aid packages designed to enhance persistence, academic achievement, and social integration (Renbarger & Long, 2019). Such varied and comprehensive supports remain insufficiently developed and implemented for autistic college students, highlighting a significant gap in institutional approaches.

Despite the encouraging outcomes reported in several studies, generalizability remains a notable concern. Many studies reviewed exhibited limited diversity within their samples due to inadequate or inconsistent demographic reporting, with several studies failing to detail participants' ethnic identities or providing only minimal information about gender diversity. Such omissions hinder the ability to ascertain whether reported intervention effects would hold true across a broader, more heterogeneous population of autistic individuals. Additionally, the reviewed literature rarely mentioned the presence or specifics of campus support programs beyond the intervention itself, and there was little discussion of how students' academic majors or departmental cultures might interact with intervention outcomes. The absence of these contextual details contrasts starkly with comprehensive interventions that have successfully supported other student groups, suggesting

interventions proven effective within narrowly defined contexts might not readily translate to settings with differing support structures or academic cultures. For instance, supports tailored specifically to nontraditional students recognize unique contextual barriers (e.g., caregiving, employment, delayed college entry; MacDonald, 2018), whereas the reviewed autism studies seldom acknowledged similarly detailed contextual elements or institution-wide supports.

Equally important is the issue of internal validity, encompassing concerns about replication, data collection, and reporting. Although multiple-baseline and multiple-probe designs were commonly employed, many studies demonstrated limited replication across participants, behaviors, or settings. This restricted replication makes it challenging to determine whether the observed improvements are robust or artifacts of the unique study conditions. Furthermore, variations in data collection methods and inconsistent reporting of detailed visual analysis procedures undermine the reliability of findings. Detailed visual analysis is critical not only for documenting the effects of interventions but also for ensuring that subtle behavioral changes are consistently detected and accurately interpreted. Additionally, systematic reviews examining postsecondary supports for students with disabilities emphasize the importance of methodological rigor, robust transition planning, and self-advocacy in effectively supporting diverse student populations in higher education (Madaus et al., 2024; Nail, 2024). Madaus et al. (2024) specifically call for more rigorous empirical research to validate existing supports, emphasizing the crucial role that social experiences and relational supports play in promoting success.

In summary, while interventions aimed at supporting autistic students in postsecondary settings show promise, particularly in enhancing social engagement and academic skills, the current body of research is limited by issues related to both generalizability and internal validity. Addressing these limitations will require diversifying participant samples, providing detailed contextual information about institutional supports and academic environments, and employing rigorous,

replicable data collection and analysis procedures. Drawing insights from robust institutional support frameworks that have already been successfully implemented for traditionally enrolled, nontraditional, and low-income students can further inform the development of future interventions. Such methodological and contextual enhancements are essential to ensure intervention outcomes are reliable and broadly applicable across the diverse landscape of higher education.

Limitations

Several limitations should be acknowledged. First, this review focused exclusively on single-case experimental designs (SCEDs), which inherently limits generalizability due to the use of small, often homogeneous samples and highly controlled intervention contexts. The absence of detailed demographic data in many studies further restricts the ability to generalize findings to broader autistic student populations. Additionally, methodological inconsistencies (e.g., minimal replication, variability in reporting visual analyses, incomplete descriptions of contextual supports) further reduce confidence in the robustness of the reported outcomes. Furthermore, I only included articles and dissertations published in English, potentially omitting relevant studies published in other languages. Lastly, because I conducted aspects of the full-text screening independently, potential bias in article selection and data extraction may have influenced findings.

Implications for Practice and Future Research

Practitioners and institutions must prioritize comprehensive and proactive support systems, reflecting successful models employed with traditionally enrolled (Lisnyj et al., 2023), nontraditional (MacDonald, 2018), and low-income students (Renbarger & Long, 2019). Campus-wide initiatives should include structured peer mentorship, enhanced social capital opportunities, tailored academic accommodations, pre-entry transition supports, and robust mental health services to address the

unique needs of autistic students (Nelson et al., 2023). Additionally, fostering faculty awareness and inclusivity training could mitigate barriers related to disclosure stigma and social isolation.

Future research should pursue methodological rigor through increased replication, detailed demographic reporting, and comprehensive contextual documentation. Future research should therefore include more diverse participant samples and explicitly consider contextual influences, including but not limited to existing support programs and academic disciplines, thereby enhancing the generalizability of findings. Furthermore, longitudinal and preventative studies explicitly designed to reduce attrition and bolster retention are urgently needed, aligning autism-specific supports with existing effective practices for other postsecondary student groups. Inclusion of studies with null or variable effects in this review is consistent with calls for methodological transparency, as recognizing and reporting non-significant outcomes helps build a balanced evidence base and highlights areas requiring methodological refinement. Such efforts will build a more robust and applicable evidence base to improve educational outcomes and overall quality of life for autistic students in higher education contexts.

Chapter 3: Meta-Analysis

The preceding chapter highlighted the growing presence of autistic students in higher education and the diverse challenges they face in campus environments. With autism prevalence continuing to rise, with recent estimates indicating that approximately 1 in 36 children are diagnosed with autism spectrum disorder (ASD; Maenner et al., 2023) and over 2% of U.S. adults are on the spectrum (Dietz et al., 2020), resulting in an increasing number of autistic young adults who are pursuing postsecondary education. However, many of these students encounter difficulties in academic, social, and daily living domains that can hinder their success in college. Colleges and universities often struggle to meet the complex needs of this population, and autistic students continue to experience lower rates of college persistence and success compared to their neurotypical peers (Petcu et al., 2021; Hamilton & Petty, 2023). This context underscores an urgent need for evidence-based interventions to support autistic college students as they navigate higher education.

In response to this need, various support programs and targeted interventions have emerged to assist autistic students in higher education settings. Recent reviews of postsecondary autism support initiatives have documented numerous programs and strategies being implemented on campuses worldwide (Nelson et al., 2023; Petcu et al., 2021). These include academic skill-building services, social skills groups, peer mentoring programs, cognitive-behavioral interventions for anxiety, technology-assisted supports, and others. However, despite the proliferation of support efforts, the effectiveness of these interventions remains under-reported. Much of the literature to date has been descriptive or qualitative, providing valuable insights into student experiences but relatively little rigorous data on which interventions actually improve student outcomes (Madaus et al., 2024). Indeed, only a small subset of studies have used experimental designs to evaluate interventions for autistic college students (Hamilton & Petty, 2023), reflecting a still-emerging evidence base. Leading scholars have called for more systematic evaluation and synthesis of

intervention research in this area (Hamilton & Petty, 2023; Nail, 2024). Building on the foundation established by the literature review, this chapter addresses the gap by introducing a meta-analysis focused on the efficacy of interventions for college students on the autism spectrum.

This meta-analysis has three primary aims: (a) to identify the types of interventions that have been implemented to support autistic college students, (b) to evaluate the overall effectiveness of these interventions as reported in single-case experimental design (SCED) studies, and (c) to examine key methodological characteristics and potential moderators that may influence intervention outcomes. In pursuing these aims, the analysis centers on empirical studies that used SCED methodologies, a common approach in educational psychology for rigorously testing interventions with individual students or small groups of students. By aggregating findings across these studies, the meta-analysis provides a quantitative assessment of how effectively various strategies improve relevant outcomes (e.g., academic performance, social interaction, mental health, adaptive skills) for students on the spectrum. Equally important, the review examines whether specific factors may impact the effectiveness of interventions. For instance, the meta-analysis examines whether differences in study design or implementation can explain the variability in results, and whether participant characteristics or intervention formats moderate the observed effectiveness. This approach enables a nuanced understanding of not only which interventions tend to be effective, but also why and under what conditions they are most successful.

To ensure a comprehensive and unbiased synthesis, data from each eligible SCED study were systematically extracted and coded by the author. This process involved building a detailed dataset that captured a range of variables related to the study context, methodology, and content. Examples of coded variables include the funding source of the study (e.g., federal grant, institutional support, unfunded), participant demographics (e.g., gender, race, age, major) that may moderate intervention outcomes, form of publication (i.e., journal article, dissertation) to determine potential

publication bias in the literature, the extent to which each study adhered to established quality indicators for SCEDs, and the critical components of each intervention (e.g., use of peer mentors, technology aids, behavioral strategies). By compiling these details, the meta-analysis is equipped to examine how methodological rigor and study features relate to outcomes. For instance, the inclusion of quality indicators allows an evaluation of whether higher-quality studies tend to report different effect sizes, and the coding of intervention components enables an analysis of which types of support yield more substantial improvements.

In summary, Chapters 1 and 2 informed a focused meta-analytic investigation into interventions for autistic college students, given the rising prevalence of autism and the perceived barriers in postsecondary settings, and has delineated the goals and scope of the analysis within the context of existing literature. The chapter now transitions from the broad overview provided here to the specifics of the meta-analysis, beginning with a description of the methods used to identify, select, and code relevant studies, followed by the aggregated results and moderator analyses. Through this systematic inquiry, the dissertation seeks to advance an evidence-based understanding of how best to support the success and well-being of autistic individuals in higher education (Maenner et al., 2023; Nelson et al., 2023; Petcu et al., 2021; Hamilton & Petty, 2023; Madaus et al., 2024; Nail, 2024).

Method

This chapter presents a meta-analysis of SCED studies that support autistic students in postsecondary settings. The studies included in this meta-analysis were identified and evaluated through the systematic review detailed in Chapter 2. The following sections summarize the methodological procedures employed in that review, which established the foundation for the current meta-analysis.

Review Methods and Search Strategy

As described in Chapter 2, my systematic review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines (Page et al., 2021). My comprehensive search strategy, detailed in Chapter 2, involved querying the ERIC and ProQuest Dissertations and Theses Global databases using an extensive Boolean string covering various terms related to neurodiversity (including autism) and postsecondary education. The search was limited to dissertations and peer-reviewed articles published from 1990 onward, aligning with significant federal transition policy shifts (Morningstar et al., 2012). An ancestral search of references in key articles was also conducted to identify additional relevant studies.

Inclusion/Exclusion Criteria

The multi-stage screening process, fully elaborated in Chapter 2 (See Figure 1 for PRISMA Flow Chart), progressively refined the article pool. For the purpose of the systematic review that informs this meta-analysis, the crucial final inclusion criteria focused on:

1. studies published in English,
2. empirical research reports (peer-reviewed articles or dissertations),
3. studies including autistic participants in postsecondary contexts, and
4. studies employing a SCED.

Key exclusions at the final stage included SCEDs not meeting minimum *NTACT:C Quality Indicators* (e.g., AB designs), qualitative studies, surveys, program evaluations, or literature reviews.

Screening Procedure and Interobserver Agreement

Initial database searches yielded 3,531 articles after duplicate removal. I screened titles and abstracts using Rayyan (rayyan.ai). As detailed in Chapter 2, IOA for this initial screening was established with a trained second reviewer on 20% of the articles (n = 706), achieving >93% agreement, with all discrepancies resolved through discussion.

Following full-text review of 51 articles based on the refined inclusion criteria, 15 studies employing SCEDs were identified. Due to logistical constraints, I conducted the initial full-text review independently, which I acknowledged as a limitation in Chapter 2. However, a trained doctoral-level colleague subsequently screened 33.3% (n = 5) of this finalized pool of 15 SCED studies, achieving 100% IOA on their inclusion for the systematic review.

Quality Assessment

I assessed the methodological quality of the 15 SCED studies identified through the systematic review, as detailed in Chapter 2, using the *National Technical Assistance Center on Transition: The Collaborative (NTACT:C) Quality Indicator Checklist for Single-Case Experimental Designs*. This instrument, also detailed in Chapter 2, evaluates aspects of internal validity, external validity, and social validity specific to SCED research. Studies were categorized as "High Quality," "Acceptable Quality," or "Did Not Meet Quality" based on adherence to 20 standardized indicators (Horner et al., 2005; Kratochwill et al., 2013), with specific criteria for initial design rigor (e.g., focusing on transition-related skills for young adults with disabilities, employing designs like multiple baseline or reversal). This quality appraisal process resulted in the exclusion of one study that failed to meet minimum NTACT:C standards, yielding a final sample of 14 studies for the present meta-analysis.

To ensure reliability in the quality assessment of the initial 15 studies, a trained doctoral-level colleague independently assessed approximately 33.3% (n = 5) of these studies using the same NTACT:C checklist. IOA at the study level averaged 91% (range: 85%-100%), and at the individual indicator level averaged 91% (range: 40%-100%). Discrepancies, primarily concerning the interpretation of practicality and naturalistic implementation indicators, were resolved through consensus discussion. The comprehensive results of this quality appraisal for each of the 15 studies, including the rationale for the excluded study, are presented in Chapter 2.

Variable Coding

To systematically characterize the studies included in this meta-analysis (which were identified and quality appraised as detailed in Chapter 2), I developed and applied a structured coding guide. This guide was meticulously informed by the domain framework of the *National Technical Assistance Center on Transition: The Collaborative* (NTACT:C) *Quality Indicator Checklist*. My approach involved examining the NTACT:C definitional terms for each domain and coding for a comprehensive range of specific variables within them to ensure thorough data capture. The key study elements coded, including those later examined as potential moderators, are detailed below.

Participant Characteristics

Comprehensive information regarding participant characteristics was coded. This included detailed demographic information such as participant age, gender, race and ethnicity, academic year in university, and declared major. Additionally, I recorded the specific diagnostic criteria used for autism, any reported co-occurring conditions, cognitive or linguistic profiles if provided by the authors, and the described processes for participant selection.

Setting Characteristics

For setting characteristics, I coded features of the postsecondary environment. This included the type of institution (e.g., 2-year, 4-year), the specific location where the intervention was implemented (e.g., classroom, resource center, community-based), and the country where the study was conducted.

Dependent Variable (DV)/Outcome Measures

Regarding dependent variables and outcome measures, I coded their operational definitions and the domain of the dependent variable (e.g., social skills, academic behaviors, adaptive behaviors). I also extracted information on the specific procedures and instruments used for

measurement, the units of measurement, the timing of these measurements, and any reported IOA or reliability data for each dependent variable.

Independent Variable (IV)/Intervention Characteristics

In relation to independent variables or intervention characteristics, I coded the operational definitions of each intervention. This included detailed descriptions of its core components, materials utilized, and specific procedures followed. Key intervention variables coded for potential moderation included the form of intervention delivery and the total time in intervention sessions in minutes. Furthermore, I recorded information on who delivered the intervention, its broader mode of delivery, and the duration, intensity, and frequency of its implementation not captured by total minutes. Any reported measures of treatment fidelity were also systematically coded.

Study Design and Procedures

For study design and procedures, I coded the specific type of SCED employed in each study, (e.g., multiple baseline, reversal design). I also extracted characteristics of the baseline phase, including its length and data stability, along with procedural details for experimental conditions, the number of participants, and the number of data points within each phase.

Findings and Evidence of Effect

Under findings and evidence of effect, I coded the authors' conclusions regarding the effectiveness of the intervention. This included any reported outcomes from visual analysis, any effect sizes or quantitative summaries of change provided in the original studies, and the reported number of demonstrations of experimental effect.

Social Validity

Regarding social validity, I coded information pertaining to the social importance of the dependent variables targeted in the studies. I also extracted data on the perceived significance of the change in outcomes, as reported by participants or other relevant stakeholders, and any descriptions

of the intervention's practicality, cost-effectiveness, or its implementation in naturalistic settings by typical intervention agents.

Overall Study Descriptors

Beyond domain-specific elements, I coded for overall study descriptors, several of which were examined as moderators. These included the NTACT:C Quality Indicator Score assigned during the quality appraisal process, the reported funding status of each study, and its publication format.

Given that I conducted this detailed study characteristic coding independently, adherence to this clearly defined coding scheme, derived from established quality indicators, was imperative for ensuring consistency and reliability across all studies. While this comprehensive coding allowed for a rich description of the study pool, elements of the specific variables highlighted above were selected for formal testing as potential moderators in this meta-analysis due to their theoretical or empirical relevance to intervention outcomes.

Data Extraction and Estimation of Effects

Data Extraction

To prepare for statistical analyses, I collaborated with a doctoral-level colleague who holds a Master's in Psychological Sciences and has research expertise in areas including student perception, motivation, and engagement. Together, we extracted raw time-series data from the graphs presented in each study using *WebPlotDigitizer*, a free software developed by Automeris. Although the tool is currently in Version 5, we used Version 4, archived as of May 15, 2024, which did not require user accounts and allowed for streamlined data extraction.

Prior to data collection, I provided structured training to my colleague. This training included instruction on how to screenshot graphs from the included articles and dissertations, convert those images to JPEG format, upload them into *WebPlotDigitizer*, calibrate the axes, identify

individual data points, extract the resulting coordinates, and compile the digitized data into a shared spreadsheet on OneDrive. The resulting dataset included all relevant conditions, including any appropriate maintenance and/or generalization phases, serving as the foundation for constructing the primary dataframe used in the subsequent meta-analytic models. This collaborative process was designed to improve accuracy, reliability, and consistency across data points extracted from complex visual representations.

Within-Case Effect Size

To quantify intervention effects, I calculated the log response ratio (LRR) for each participant, comparing their mean performance during intervention (M^I) phases to baseline (M^B) phases. This calculation was performed using the *Single-Case Effect Size Calculator* (Version 0.7.3). The LRR, calculated as $LRR=\ln\left(\frac{M_I}{M_B}\right)$, measures the proportional change in outcome levels and is particularly well-suited for ratio-scale data where zero indicates a true absence of the behavior (Pustejovsky, 2018). The calculator (Pustejovsky et al., 2024) is available online at https://jepusto.shinyapps.io/SCD-effect-sizes/.

Positive LRR_i values (i.e., Log Response Ratio - Increasing) indicate therapeutic trends, while negative LRR_d values (i.e., Log Response Ratio - Decreasing) can reflect contra-therapeutic trends or desired decreases, depending on the outcome. For frequency or rate outcomes, LRR_i and LRR_d typically share magnitude but differ in sign; however, for percentage or proportion outcomes, their interpretation and calculation can diverge in both respects (Pustejovsky, 2018).

Meta-Analytic Model

After computing the LRR effect size for each AB phase contrast across studies, I built the meta-analytic model. Because each study had multiple, dependent effect sizes I built a random effects, multilevel model using robust variance estimation to account for dependent effect sizes. I evaluated heterogeneity by reporting τ, τ^2, and a 95% prediction interval. I investigated the potential

impact of publication bias through the visual analysis of funnel plots and Egger regression tests. Last, due to the limited sample of studies included (n = 14), building a meta-regression model for moderator analysis that included all variables of interest was not possible. Thus, I conducted univariate moderator analyses to explore the potential explanatory power of specific variables.

Results

Summary of Effects

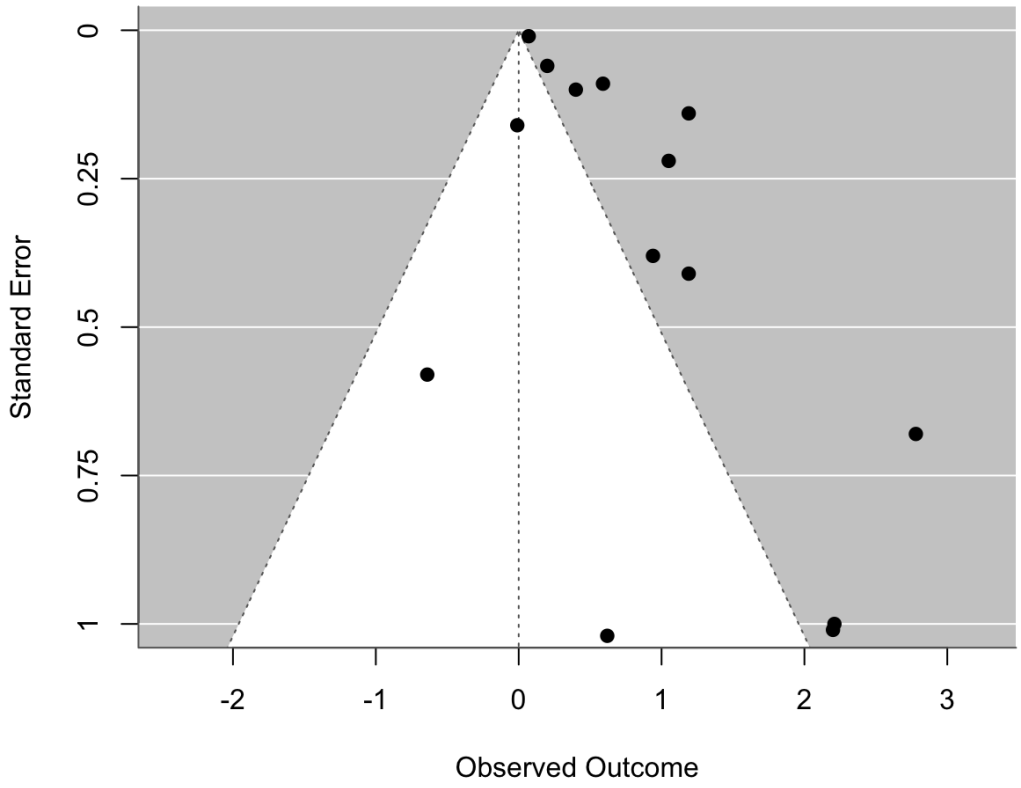
In the two-level random-effects model, the synthesized intervention effect was LRR = 0.83 (SE = 0.20; 95% CI [0.43, 1.23], $p < .0001$; R Markdown available at https://doi.org/10.17605/OSF.IO/V6NTC, denoting a moderate-to-large and statistically significant benefit. However, the 95% prediction interval (-1.20, 2.86) was correspondingly broad, suggesting that the true effect in a novel study may range from a slight detriment to a substantial gain. Heterogeneity was pronounced, with estimated between-study variability ($\tau_1 = 0.61$) and within-study variability ($\tau_2 = 0.84$), indicating considerable dispersion at both hierarchical levels. Cochran's Q-statistic (Q = 1626.98, $p < .0001$) corroborated that residual heterogeneity exceeded sampling fluctuation, thereby justifying subsequent moderator analyses to elucidate sources of effect-size variation.

Publication Bias

The funnel plot reports the standard error of each LRR against its estimate, showing significant asymmetry. In the full-effects funnel (n = 109; See Appendix A: Figure 2), the cloud of low-precision estimates fans out toward both tails, with an increased number of extreme negative LRR values, suggesting that small, individual comparisons may both inflate positive findings and occasionally produce unusually large negative outcomes. By contrast, when those same data are aggregated into study-level (n = 14; See Figure 2) summaries, the corresponding funnel plot becomes markedly more compact and symmetrically distributed around the null line, with far fewer deviant

data points. Egger's regression test for funnel plot asymmetry provided statistical confirmation of this observed asymmetry ($z = 5.36$, $p < .0001$). The intercept (bias coefficient) from this test was estimated at b = 0.06 (95% CI [–0.21, 0.33]). Combined, these patterns suggest the potential presence of small-study effects or publication bias in the meta-analytic literature, particularly at the individual effect-size level.

Figure 2. Funnel Plot of Effects



Moderator Analyses

Given the significant heterogeneity identified in the overall analysis of intervention effects, derived from 109 effect sizes across the 14 included SCED studies, I conducted moderator analyses to explore potential sources of this variability. Moderator variables were selected a priori based on

their theoretical relevance to intervention outcomes for autistic students in postsecondary settings, findings suggested by previous literature, and the availability of consistent reporting of these characteristics across the included studies. The detailed results for each moderator tested, including the number of studies (N), number of effect sizes (ES), and specific effect size estimates for relevant subgroups, are presented comprehensively in Table 4 and elaborated upon in the subsequent sections of this chapter.

Table 4. Multivariate Moderator Analyses

Moderator	N (ES)	LRR (SE)	95% CI	p-value
NTACT Quality Standards				
Does Not Meet	3(51)	0.32 (0.37)	-0.41, 1.04	.3854
Meets	11(58)	0.99 (0.22)	0.56, 1.42	<.0001
Gender				
Female	7 (23)	0.86 (0.28)	0.30, 1.41	.0028
Male	14 (86)	0.83 (0.21)	0.41, 1.24	.0001
Race				
BIPOC	5 (8)	1.02 (0.40)	0.23, 1.82	.012
White	8 (31)	0.98 (0.29)	0.41, 1.55	.0009
NR	6 (70)	0.64 (0.31)	0.03, 1.24	.04
Age				
Continuous (Reported)	13 (85)	0.007 (0.001)	-0.01, 0.03	.4764
Year				
1 (Freshman)	7 (30)	0.61 (0.29)	0.04, 1.17	.0357
2 (Sophomore)	2 (15)	0.56 (0.40)	-0.23, 1.35	.1649
3 (Junior)	3 (7)	0.83 (0.43)	-0.03, 1.68	.0596
4 (Senior)	3 (7)	1.194 (0.45)	0.31, 2.08	.0089
NR	8 (50)	0.97 (0.26)	0.45, 1.48	.0003
Major				
Art	2 (4)	-0.01 (0.60)	-1.2, 1.17	.9823
Social Science	4 (7)	1.42 (0.43)	0.58, 2.27	.0011
STEM	7 (15)	0.81 (0.31)	0.20, 1.42	.0093
Undeclared	3 (5)	1.05 (0.49)	0.08, 2.03	.0342
NR	9 (78)	0.75 (0.22)	0.31, 1.19	.0011
Publication				
Dissertation	3 (30)	0.26 (0.39)	-0.51, 1.03	.5101
Journal	11 (79)	0.98 (0.21)	0.56, 1.41	<.0001
Funding				
Yes	4 (20)	1.08 (0.39)	0.29, 1.86	.0073
No	4 (27)	1.26 (0.37)	0.52, 2.00	.001
NA (Dissertation)	3 (30)	0.27 (0.42)	-0.56, 1.09	.5244
NR	3 (32)	0.58 (0.42)	-0.25, 1.39	.1691
Dependent Variable (Group)				
Academic	4 (14)	0.76 (0.43)	-0.09, 1.61	.0823
Adaptive Behavior	2 (4)	1.05 (0.75)	-0.44, 2.55	.1636
Social	8 (91)	0.86 (0.28)	0.30, 1.42	.0028
Independent Variable (Group)				
Academic Instruction	2 (5)	0.54 (0.57)	-0.59, 1.68	.3446
Animal Therapy	1 (24)	-0.03 (0.63)	-01.3, 1.22	.9612
Meta-Cognition	1 (24)	0.47 (0.65)	-0.82, 1.77	.466
Social/Peer Mediation	3 (15)	1.65 (0.46)	0.74, 2.56	.0005
Technology	3 (10)	1.09 (0.49)	0.12, 2.07	.0286
Video-Based	4 (31)	0.69 (0.35)	-0.001, 1.37	.0505
Total Minutes				
Continuous (Recorded)	11 (81)	0.0005 (0.0005)	-0.0004, 0.0015	.2378

Note. N = number of studies; (ES) = number of effect sizes; (SE) = standard error; 95% CI = 95% confidence interval; NR = not reported.

NTACT Quality Indicators

The omnibus test indicated that NTACT quality significantly moderated effect sizes ($F = 10.67$, $p < .0001$). Studies meeting the minimal NTACT standards produced a large, significant effect (LRR = 0.99, SE = 0.22; 95 % CI [0.56, 1.43], $p < .0001$), while studies that did not meet these standards showed a smaller, non-significant effect (LRR = 0.32, SE = 0.37; 95 % CI [-0.41, 1.05], $p = .385$). Despite accounting for quality, substantial heterogeneity remained ($Q = 1539.10$, $p < .0001$), with $\tau_1 = 0.55$ at the study level and $\tau_2 = 0.84$ at the effect level, underscoring that additional factors beyond reported methodological rigor contribute to variability in outcomes.

Gender

The omnibus test for gender indicated significant moderation of intervention outcomes ($F = 8.36$, $p = .0004$), indicating that the effect of the intervention differed by gender. Examining subgroup effects, participants identifying as female showed a large, significant benefit (LRR = 0.86, SE = 0.28; 95 % CI [0.30, 1.41], $p = .0028$), while those identifying as male exhibited a similarly robust effect (LRR = 0.83, SE = 0.21; 95 % CI [0.41, 1.24], $p = .0001$). Despite accounting for gender, substantial residual heterogeneity remained ($Q = 1574.6$, $p < .0001$), with variance components at both the study level ($\tau_1 = 0.62$) and the effect level ($\tau_2 = 0.84$), suggesting that additional moderators must be investigated to account for further variability.

Race/Ethnicity

Examining race and ethnicity ($n = 109$), the omnibus test confirmed significant differences across groups ($F = 3.63$, $p = .0046$). Disaggregated estimates revealed that Latinx participants experienced a large, significant benefit (LRR = 1.14, SE = 0.56; 95 % CI [0.03, 2.25], $p = .044$), and White participants showed a similarly strong effect (LRR = 0.97, SE = 0.28; 95 % CI [0.41, 1.54], $p = .0009$). Although Asian (LRR = 0.76, $p = .20$) and Middle Eastern (LRR = 1.50, $p = .16$) subgroups exhibited positive trends, these did not reach statistical significance. Notably, studies

lacking demographic reporting (i.e., "NR;" n = 70) still produced a conventional effect (LRR = 0.63, p = .038), highlighting the value of transparent participant-characteristic documentation. Despite accounting for these categories, unexplained variability remained high (Q = 1351.45, p < .0001), with τ_1 = 0.61 at the study level and τ_2 = 0.85 at the effect level, indicating that further moderators should be explored.

Aggregating participants into three broad categories (i.e., BIPOC, White, and Not Reported) produced a significant omnibus test (F = 5.76, p = .0011), indicating that these groupings meaningfully differentiate intervention effects. BIPOC participants showed the largest effect (LRR = 1.03, SE = 0.40; 95 % CI [0.23, 1.82], p = .012), White participants demonstrated a similarly strong benefit (LRR = 0.98, SE = 0.29; 95 % CI [0.41, 1.55], p = .0009), and the NR category remained significant (LRR = 0.64, SE = 0.31; 95 % CI [0.03, 1.24], p = .040). Despite a more parsimonious fit (τ_1 = 0.62 at the study level; τ_2 = 0.84 at the effect level), residual heterogeneity persisted (Q = 1354.96, p < .0001). Notably, aggregating into BIPOC reduced its SE from approximately 0.56–1.06 in disaggregated subgroups to 0.40, improving estimate precision while acknowledging that such pooling may obscure nuances among specific racial and/or ethnic groups.

Age

Age was initially examined as a categorical moderator across three bucketed groups (i.e., 18-24, 24+, and NR), with 18-24 selected in alignment with reports indicating that 64% of students complete a bachelor's degree within six years (National Center for Education Statistics, 2022). The omnibus test of age in this categorical analysis was significant, F = 5.5 (p = .0015), indicating that effect sizes differed across age groups. Participants aged 18–24 showed a moderate, significant benefit (LRR = 0.82, SE = 0.24; 95 % CI [0.35, 1.28], p = .0008), while those 24+ experienced an even larger gain (LRR = 1.13, SE = 0.36; 95 % CI [0.41, 1.84], p = .0023). Effects for the NR group were smaller and non-significant (LRR = 0.48, SE = 0.72; 95 % CI [–0.93, 1.89], p = .50). Despite

this categorical grouping, residual heterogeneity remained high (Q = 1463.89, $p < .0001$), with τ_1 = 0.66 at the study level and τ_2 = 0.83 at the effect level, indicating substantial unexplained variability and the need to explore additional moderators.

Testing age as a continuous predictor (n = 85), the linear slope did not reach significance (F = 0.51, p = .48). While the model intercept was significant (LRR = 0.72, SE = 0.33; 95 % CI [0.07, 1.37], p = .031), the slope coefficient (b = 0.007, SE = 0.009; 95 % CI [–0.01, 0.03]) did not differ from zero, indicating no linear trend in effect size across age. Residual heterogeneity remained elevated (Q = 746.82, $p < .0001$), with τ_1 = 0.86 at the study level and τ_2 = 0.24 at the effect level, suggesting that age, when treated continuously, does not account for the substantial unexplained variability.

Academic Year

Academic year significantly moderated intervention effects (F = 3.76, p = .0036). First-year students experienced a moderate, significant benefit (LRR = 0.61, SE = 0.29; 95 % CI [0.04, 1.17], p = .036), while second-year effects were smaller and non-significant (LRR = 0.56, SE = 0.40; 95 % CI [–0.23, 1.35], p = .165). Third-year students showed a positive trend that approached significance (LRR = 0.83, SE = 0.43; 95 % CI [–0.03, 1.68], p = .0596), with fourth-year participants demonstrating the largest, significant gain (LRR = 1.19, SE = 0.45; 95 % CI [0.31, 2.08], p = .0089). Although the NR group showed a robust, significant effect (LRR = 0.97, SE = 0.26; 95 % CI [0.45, 1.48], p = .0003), the results underscore how missing demographic data can mask actual subgroup differences. Residual heterogeneity remained high (Q = 1544.30, $p < .0001$), with τ_1 = 0.60 at the study level and τ_2 = 0.85 at the effect level, indicating that additional moderators likely contribute to the variability in outcomes.

Academic Major

Academic major significantly moderated intervention outcomes ($F = 4.83$, $p = .0005$), demonstrating variability across domains. Art majors showed no observable benefit (LRR = –0.01, SE = 0.60; 95 % CI [–1.20, 1.17], $p = .98$), whereas social-science students experienced the most considerable effect (LRR = 1.42, SE = 0.43; 95 % CI [0.58, 2.27], $p = .0011$). Undeclared majors also saw a substantial gain (LRR = 1.06, SE = 0.49; 95 % CI [0.08, 2.04], $p = .0342$), and STEM students benefited significantly (LRR = 0.81, SE = 0.31; 95 % CI [0.20, 1.42], $p = .0093$). Participants with unreported majors ("NR") still demonstrated a moderate, significant effect (LRR = 0.75, SE = 0.22; 95% CI [0.31, 1.19], $p = .0011$), highlighting the importance of transparency in demographics. Despite this moderation, residual heterogeneity remained high ($Q = 1541.93$, $p < .0001$), with $\tau_1 = 0.55$ at the study level and $\tau_2 = 0.84$ at the effect level, indicating that further moderators are needed to explain the remaining variability.

Publication Type

Publication type significantly moderated effect sizes ($F = 10.86$, $p < .0001$), with dissertations producing a small, non-significant effect (LRR = 0.03, SE = 0.39; 95 % CI [–0.51, 1.03], $p = .51$), whereas journal articles yielded a large, significant benefit (LRR = 0.98, SE = 0.21; 95 % CI [0.56, 1.40], $p < .0001$). Despite accounting for publication format, residual heterogeneity remained high ($Q = 1624.76$, $p < .0001$), with $\tau_1 = 0.54$ at the study level and $\tau_2 = 0.84$ at the effect level, indicating additional factors beyond publication type drive variability in observed outcomes.

Funding

Funding status significantly moderated effect sizes ($F = 5.30$, $p = .0006$). Studies without funding ("No") produced the largest effects (LRR = 1.26, SE = 0.37; 95 % CI [0.52, 2.00], $p = .001$), followed by funded studies ("Yes"; LRR = 1.08, SE = 0.39; 95 % CI [0.30, 1.86], $p = .0073$). Effects were smaller and non-significant when funding status was not reported ("NR"; LRR = 0.58, SE =

0.42; 95 % CI [–0.25, 1.40], p = .169) or in dissertation studies where funding was not applicable ("NA"; LRR = 0.27, SE = 0.42; 95 % CI [–0.56, 1.09], p = .524). Despite this moderation, residual heterogeneity remained high (Q = 1452.35, p < .0001), with τ_1 = 0.60 at the study level and τ_2 = 0.84 at the effect level, indicating that additional factors beyond funding status contribute to variability in outcomes.

Outcome Domain

Targeted outcome domains significantly moderated effect sizes (F = 4.80, p = .0036), with social outcomes exhibiting the most consistent benefit (LRR = 0.86, SE = 0.28; 95% CI [0.30, 1.42], p = .0028), indicating a reliable improvement on social measures. Academic outcomes showed a moderate, non-significant effect (LRR = 0.76, SE = 0.43; 95 % CI [–0.10, 1.61], p = .0823), and adaptive behavior outcomes produced the largest point estimate but did not reach significance (LRR = 1.06, SE = 0.75; 95 % CI [–0.44, 2.55], p = .16), reflecting considerable imprecision. Despite these domain distinctions, residual heterogeneity remained high (Q = 1590.09, p < .0001), with τ_1 = 0.71 at the study level and τ_2 = 0.84 at the effect level, suggesting further moderators are needed to explain the variability in outcomes.

Intervention Delivery

Intervention delivery method significantly moderated effect sizes (F = 3.88, p = .0016). Structured social planning and/or peer mediation produced the largest benefit (LRR = 1.65, SE = 0.46; 95 % CI [0.74, 2.56], p = .0005), and technology-based approaches also yielded a robust, significant effect (LRR = 1.09, SE = 0.49; 95 % CI [0.12, 2.07], p = .0286). Video-based interventions approached significance (LRR = 0.69, SE = 0.35; 95 % CI [–0.00, 1.37], p = .0505), suggesting promise but greater variability. Academic instruction (LRR = 0.54, p = .345), animal-assisted methods (LRR = –0.03, p = .961), and meta-cognitive strategies (LRR = 0.48, p = .466) did not differ significantly from zero. Despite these delivery-mode distinctions, residual heterogeneity

remained high (Q = 1390.67, $p < .0001$), with $\tau_1 = 0.60$ at the study level and $\tau_2 = 0.83$ at the effect level, indicating that other moderators should be explored to understand variance further.

Intervention Duration

The test of total intervention minutes as a continuous moderator was not significant (F = 1.42, $p = .2378$). The model intercept, representing the average effect when minutes = 0, was significant (LRR = 0.62, SE = 0.30; 95 % CI [0.03, 1.22], $p = .0397$), but the slope did not differ from zero, indicating no linear relationship between duration and effect size. Residual heterogeneity remained high (Q = 726.00, $p < .0001$), with $\tau_1 = 0.77$ at the study level and $\tau_2 = 0.26$ at the effect level, suggesting that intervention length alone does not explain the observed variability.

Discussion

In this meta-analysis, I synthesized 109 effect sizes across 14 SCED studies examining interventions for autistic college students. Overall, interventions yielded a moderate-to-large benefit (g = 0.83, SE = 0.20; 95% CI [0.43, 1.23]; $p < .0001$), indicating that, on average, targeted supports significantly improve outcomes for this population. However, considerable heterogeneity (Q = 1626.98; $p < .0001$) at the study level ($\tau_1 = 0.61$) and the effect level ($\tau_2 = 0.84$) suggests variability in efficacy across contexts and intervention types.

Examination of Substantial, Significant Effects

Among delivery methods, structured social planning and peer-mediated supports produced the largest effect (LRR = 1.65, SE = 0.46; 95% CI [0.74, 2.56], $p = .0005$), aligning with calls for peer-based interventions that leverage social networks to enhance belonging and engagement (Nelson et al., 2023). Although the pooled effect was significant, the confidence interval is quite wide, indicating that while we can be confident in the direction and general magnitude of the effect, the estimate remains imprecise, potentially limited by small sample sizes and study-level variability. Technology-assisted interventions similarly demonstrated a robust benefit (LRR = 1.09, SE = 0.49;

95% CI [0.12, 2.07], p = .0286), underscoring the promise of digital platforms to scaffold executive function and communication skills for autistic students (Hamilton & Petty, 2023). Yet, as with structured social planning and peer mediation supports, the broad confidence interval suggests imprecise measurement, likely due to small samples or between-study heterogeneity.

In terms of outcome domains, social measures yielded a consistent, significant improvement (LRR = 0.86, SE = 0.28; 95% CI [0.30, 1.42], p = .0028), indicating that interventions may be particularly effective at addressing the interpersonal challenges that often hinder persistence and well-being in higher education (Madaus et al., 2024; Nelson et al., 2023; Petcu et al., 2021). Methodological rigor also mattered, with studies meeting NTACT Quality Indicators reporting substantial gains (LRR = 0.99, SE = 0.21; 95% CI [0.56, 1.42], p < .0001), whereas those not meeting minimum standards did not yield significant effects (LRR = 0.32, SE = 0.37; 95% CI [–0.41, 1.04], p = .385), highlighting the importance of adherence to SCED standards for reliable, replicable findings (Kratochwill et al., 2023).

Positive-Negative Effects and Potential Implications

Although most intervention typed produced positive effects, a few methods showed negligible or negative trends. Animal-assisted (LRR = -0.03, p = .96), academic instruction (LRR = 0.54, p = .34), and metacognitive (LRR = 0.48, p = .47) interventions failed to reach significance. These findings suggest that passive or unidimensional approaches may be insufficient to address the multifaceted challenges autistic students face, particularly social and emotional aspects that are crucial for retention and success (Petcu et al., 2021; Hamilton & Petty, 2023). The presence of a small number of negative or non-significant effects also highlights the risk of "one size fits all" programming and underscores the need for individualized, comprehensive support.

Limitations

There are several limitations I recognize in this meta-analysis. First, I conducted segments (i.e., full-text screening) and early coding independently, although 30% of the initial pool and 30% of the final pool of texts were dual reviewed, which may have introduced selection bias. Second, when the baseline means equaled zero, the LRR could not be defined, resulting in excluded cases that likely omitted very low-frequency behaviors and may have biased effect estimates (Pustejovsky, 2018). Similarly, studies that reported fewer than three data points per phase cannot be calculated, reducing the comprehensiveness of the meta-analysis and underrepresenting interventions with sparse measurement schedules.

Additionally, by limiting the review to English-language SCEDs, the small, homogeneous samples under controlled conditions may constrain the variability in real-world settings. Despite using NTACT Quality Indicators, substantial residual heterogeneity suggests that unmeasured factors (e.g., implementation fidelity, contextual supports) potentially affected outcomes. Finally, I observed small-study effects and funnel-plot asymmetry (Egger's $z = 5.36$, $p < .0001$), indicating potential publication bias.

Implications for Practice and Future Research

The robust, moderate-to-large effects observed for structured social planning and peer mediation interventions (LRR = 1.65, $p = .0005$) and technology-based interventions (LRR = 1.09, $p = .0286$) underscore the value of extending these approaches across the transition from secondary to postsecondary settings. High schools and transition-aligned programs should integrate structured peer-support components, including but not limited to peer mentoring, social skills groups, and collaborative learning strategies to build social capital and self-advocacy skills before entering postsecondary education and/or employment (Madaus et al., 2024; Nail, 2024; Nelson et al., 2023; Petcu et al., 2021). Embedding technology tools (e.g., wearable technology, organizational

application, virtual coaching platforms) into transition programs for secondary-aged students, educators can scaffold executive functioning and communication, preparing students to self-manage the complex academic and social demands in postsecondary environments (Black et al., 2020; Hamilton & Petty, 2023; Wright et al., 2022)

Postsecondary institutions should partner with secondary schools, particularly feeder districts, to create continuity in supports, sharing relevant data on students' strengths and needs, co-designing bridge programs, and offering summer orientations that include peer-mediated social activities and technology training (Renbarger & Long, 2019). Furthermore, implementing Universal Design for Learning (UDL) frameworks can further enhance accessibility by providing multiple means of engagement, representation, and expression, thereby reducing barriers for neurodiverse learners (Hamilton & Petty, 2023; Nail, 2024).

Despite the growing body of SCED literature examining autistic postsecondary students, this meta-analysis revealed substantial heterogeneity and significant evidence of publication bias, indicating gaps in both the breadth and representativeness of available studies. Future meta-analyses should broaden inclusion criteria to encompass interventions targeting autistic individuals in diverse contexts, including but not limited to community college, vocational education, online learning environments, and on-the-job training supports to capture variability and nuance in program models and demographics.

Moreover, expanding beyond exclusively autistic samples to include other neurodivergent populations (e.g., ADHD) who may encounter similar barriers to full inclusion and participation in postsecondary contexts will allow the examination of transdiagnostic intervention components and facilitate comparisons of intervention efficacy across neurodiversity profiles. Researchers should pursue detailed reporting of participant demographics, including beyond surface-level factors (e.g., comorbidity, major, socio-economic status), specifically in SCEDs, which require thorough

reporting to facilitate replication and further generalizability. Finally, integrating grey literature, such as unpublished dissertations, evaluation reports, and null or negative effect studies, can provide a more comprehensive understanding of effects.

Chapter 4: Single-Case Experimental Design

Postsecondary enrollment among autistic individuals is on the rise, with many pursuing traditional college degrees (Anderson et al., 2017; Dwyer et al., 2023; Schembri-Mutch et al., 2023). Despite this increase, existing literature indicates that autistic individuals enroll in postsecondary education at significantly lower rates than their allistic peers and those with speech or language impairments, learning disabilities, or cognitive disabilities. Specifically, enrollment rates for autistic students have remained at 44% across all levels of postsecondary education, a figure that is lower than nearly all other disability classifications, except for multiple disabilities (32.8%) and specific categories of intellectual and developmental disabilities (28.7%; Anderson et al., 2017; Newman et al., 2011; Schembri-Mutch et al., 2023).

Moreover, graduation rates for the 44% of autistic students who enroll in postsecondary education are significantly lower (39%) compared to their allistic peers (59%) and peers with other disability classifications (50%; Petcu et al., 2021). Research on the specific factors contributing to why autistic students may leave institutions of higher education before degree completion is limited and predominantly qualitative. Existing studies suggest that these factors are often linked to challenges to mental well-being, driven by a complex interplay of elements, including but not limited to (a) the "hidden curriculum" of university life, encompassing unspoken norms related to behavior, study habits, and social interactions; (b) negative perceptions of education stemming from unsupportive primary and secondary school environments; (c) challenges in establishing a sense of belonging, often intensified by overwhelming social situations that can lead to bullying or feelings of isolation; and (d) the phenomenon of masking, where autistic individuals suppress their neurodivergent responses to “fit in,” resulting in exhaustion, dissociation, psychological distress, and an increased likelihood of leaving before completing their degrees (Hamilton & Petty, 2023; Petcu et al., 2021).

While existing literature often highlights the barriers autistic degree-seeking students face in postsecondary education, there is a notable lack of research focused on identifying solutions to these challenges. This study seeks to fill that gap by documenting the social engagement and interaction styles of traditionally enrolled autistic undergraduates to understand their connection to campus life and major-specific programming. "Traditionally enrolled" refers to students actively participating in the general education curriculum on campus, whether or not they receive support services and accommodations from the institution's Accessibility and Disability Resource Center (ADRC). Specifically, this study will gather input from first- and second-year autistic students who are entering science, technology, engineering, or mathematics (STEM) or STEM-adjacent fields (e.g., accounting, broadcasting) and are involved in an on-campus support program designed for degree-seeking autistic students.

Study Purpose

This study aims to evaluate whether the PEERS® for Young Adults (PEERS®-YA) social learning intervention, designed for autistic youth and young adults, enhances social engagement on university campuses, particularly among students enrolled in STEM and STEM-adjacent fields. Additionally, the study will investigate whether increased social engagement correlates with a stronger sense of belonging, both on campus and within specific academic programs. Existing literature suggests that traditionally enrolled, autistic degree-seeking students experience higher departure rates before degree completion compared to their allistic peers and peers with other disability classifications, with challenges related to social interaction and establishing a sense of belonging often identified as critical factors impacting degree completion. This study will explore whether there is a functional relation between enhanced social behaviors and a student's sense of belonging by addressing the following research questions:

a. Is there a functional relation between the PEERS® for Young Adults social learning intervention and increased social engagement among traditionally enrolled autistic students in STEM or STEM-adjacent fields?

b. Is there a functional relation between increased social engagement on campus and within an academic program and a student's sense of belonging on campus and within their major?

c. Is there a functional relation between active social engagement and a student's overall feelings of positivity toward their college experience?

Methods

This study was registered with the Open Science Framework (OSF; Beaty, 2025; https://doi.org/10.17605/OSF.IO/V6NTC). I conducted this study as part of my final dissertation, which was supported by an Office of Special Education Programs (OSEP) grant. Additional support was provided through my involvement with an on-campus program that assists autistic students, with priority participation given to individuals already enrolled in the program.

Participants

This study focused on first- and second-year autistic university students who are traditionally enrolled in STEM fields. It aims to address the well-documented challenge of higher rates of departure before degree completion among this population in postsecondary contexts. Given the scope of this doctoral dissertation, participants were recruited through convenience sampling from an on-campus support program specifically designed for autistic undergraduates pursuing STEM disciplines.

Although the pilot cohort exhibited limited diversity, the support program's second year on campus experienced a notable increase in participant diversity, including greater representation of various racial backgrounds, gender identities, and sexual orientations. This enhanced diversity

enriched the participant pool and improved the generalizability of the study's findings. However, any remaining limitations due to participant homogeneity are explored in the analysis.

PEERS®-YA recommends forming collaborative groups consisting of approximately seven to ten young adults (aged 18–30), without restrictions on gender distribution. Prior observations suggest that forming groups exclusively composed of autistic individuals may enhance comfort and participation among group members (Laugeson, 2017). In alignment with these recommendations and considering the available resources, I intended to adhere to the PEERS®-YA framework by (a) prioritizing students enrolled in the on-campus support program, (b) forming cohorts of four to seven students from this program, and (c) limiting the intervention to first- or second-year autistic degree-seeking students, aged 18 to 30. All students enrolled in the on-campus support program were offered the intervention; however, only those who provided informed consent were included in the data collection process. Given few participants consented to participate in the research, the primary intervention was modified to allow one-on-one sessions rather than a group focus, with limitations of this approach discussed in that section.

Participant Descriptions

Chase is a 20-year-old, second-year heterosexual male (he/him) majoring in Meteorology. He is autistic, diagnosed at age nine, and has co-occurring ADHD. He is registered with the university's ADRC, receiving extended-time exam accommodations in a distraction-free environment. He also participates in an on-campus support program for autistic undergraduates. Chase attended weekly meetings and proctored study halls and was invited to 18 workshops on topics, including but not limited to (a) goal-setting, (b) time management, (c) executive functioning, (d) academic skills, and (e) social engagement. While maintaining near-perfect attendance at meetings and study halls, he did not attend any of these 18 workshops.

Outside of these services, Chase was active in the Meteorology Club and interned with the University Broadcasting Network, with the intention of combining his interests in atmospheric science and broadcasting into a future career. He lived on campus during his first year before relocating off campus and working part-time for his family. Recruitment occurred via the support program's invitations, which included: (a) a program-wide email, (b) an in-person announcement by the program director, and (c) direct contact from the research team. I verified Chase's diagnostic status through his initial program application and offered him the opportunity to consult an independent party with questions. He provided informed consent under an IRB-approved protocol, re-consenting after study remuneration was increased.

Alyx, a second-year homosexual genderqueer male (they/them) majoring in Computer Science, is autistic (diagnosed at age sixteen) and has co-occurring ADHD, anxiety, and depression. Registered with the university's ADRC, Alyx receives accommodations such as extended time for exams in a distraction-free setting and the option to request extensions on assignments.

Alyx participates in the on-campus support program for autistic undergraduates, regularly attending weekly meetings and proctored study halls. From the 18 workshops offered on topics previously mentioned (e.g., goal-setting, academic skills), Alyx attended seven.
Outside of this program, Alyx maintained active involvement in the Queer Student Association and joined a bi-monthly tabletop roleplaying group with peers. As a local resident, they have lived off campus since matriculation.

Alyx was recruited through the same program-wide invitations and direct contact from the research team, as previously detailed. I verified their diagnostic status via their initial program application, offered the chance to consult with an independent third party regarding any questions, and obtained informed consent under an IRB-approved protocol.

Table 5. Participant Demographics

Participant	Gender Identity	Age	Sexual Identity	Major	Year	Diagnosis	Comorbidity
- Chase	Male	20	Heterosexual	Meteorology	2	ASD	ADHD
- Alyx	Genderqueer Male	20	Homosexual	Computer Science	2	ASD	ADHD; Anxiety; Depression

Setting

The study was conducted at a large midwestern university in a suburban community serving suburban, urban, and rural populations. All intervention sessions were held on the university campus in the conference room or writing center where the on-campus support program for autistic students is housed. This setting provided a familiar and accessible environment for participants, facilitating their engagement in the intervention.

Social activities were conducted within each participant's natural environment, such as their residence, university facilities, or community settings. This approach enabled the application of acquired skills in real-world contexts, thereby enhancing the ecological validity of the intervention. I monitored self-reported data on social engagement levels and overall sense of belonging throughout the study to assess the intervention's effectiveness. While most of the data collection occurred in these natural environments, the on-campus support program included structured social events (e.g., game nights, dinner gatherings), which provided participants with opportunities to refine their skills in controlled yet realistic settings. These events took place in university facilities, various community spots, and other settings suitable for social interaction.

Materials

This study relied on experience sampling methodology (ESM) to collect self-reported data, a method that has grown significantly due to more accessible delivery options like smartphones and smartwatches (Atz, 2013; Fritz et al., 2024; Mölsä et al., 2022; Verhagen et al., 2016; von Klipstein et

al., 2023). The study was conducted using m-Path (https://m-path.io/), an online platform designed for practitioners to extend treatment beyond clinical settings and explicitly developed for real-life assessment and intervention. While I will use the platform's dashboard via a traditional internet connection on a university-approved laptop, participants will install the m-Path application on their personal electronic devices (e.g., smartphone, smartwatch). This setup allows them to receive push notifications and quickly complete surveys and daily prompts throughout all study phases, including pre- and post-intervention data collection.

Time Series Data. ESM involves repeated measures over time, potentially placing a significant burden on participants. To promote compliance and minimize missing data, it is recommended that measures be completable within one to three minutes (Fritz et al., 2024; Verhagen et al., 2016). To assess day-to-day experiences of belonging, I selected four primary questions, with a fifth write-in option. These questions were: (1) At this moment, how strong is your sense of connectedness or belonging with others? Responses were measured on a slider, ranging from 0 (i.e., "Not Connected") to 100 (i.e., "Extremely Connected"). (2) Were you interacting with others during this time? Responses included "Yes" or "No," leading into the following question. (3) If you answered "Yes" to the previous question, who were you with? (4) How would you rate the overall positivity of your social interactions in the past hour? Responses were similarly measured on a slider, ranging from 0 (i.e., "Negative") to 100 (i.e., "Positive"). (5) If you have any additional thoughts or feelings you'd like to share about your sense of belonging today, please write them below. The full survey instrument, including images of the m-Path interface, can be found at https://doi.org/10.17605/OSF.IO/V6NTC.

Pre-Post Assessment. To monitor participants' growth in social skills knowledge, I administered the *Test of Young Adult Social Skills Knowledge* (TYASSK), a 30-item criterion-referenced test specifically developed for PEERS®-YA. The TYASSK served as both a pre- and post-

intervention measure, with higher scores indicating greater knowledge of social skills. Previous studies utilizing the TYASSK to assess treatment outcomes have generally reported positive results for autistic individuals who received the intervention, yet formal psychometric properties for the TYASSK have not been examined for PEERS-YA (Laugeson, 2017).

While PEERS®-YA has a rich research history, and other assessments like the *Social Responsiveness Scale—Second Edition*, the *Social Skills Improvement System*, and the *Social and Emotional Loneliness Scale for Adults* could evaluate social interactions, budgetary constraints limit the use of these instruments in this study. Additionally, PEERS®-YA typically employs the *Quality of Socialization Questionnaire* to assess the frequency of social events and the level of conflict within them. However, as this study's data collection process emphasizes "in-the-moment" feelings, this reduces reliance on this particular measure.

To monitor participants' growth in their sense of belonging, I used a modified version of the *Social Connectedness Scale–Revised* (SCSR). Originally developed for youth and young adults, the SCSR is a 20-item self-report instrument measuring the degree to which individuals feel connected to others in their social environment. Each item is rated on a Likert scale from "Strongly Disagree" to "Strongly Agree." Negatively worded items are reverse-coded, and higher total scores indicate a stronger sense of belonging. The SCSR has demonstrated high reliability (internal consistency $\alpha >$ .92) and strong validity across concurrent, construct, convergent, and discriminant domains (Lee & Robbins, 1995; Sabitelli & Anderson, 2005). For this study, the SCSR was adapted to focus on the college experience through minor item modifications. For instance, "I feel disconnected from the world around me" was changed to "I feel disconnected from the university." Additionally, the number of items was reduced to 10, and the standard Likert scale was replaced with a slider (0 to 100) anchored by "Disagree" and "Agree." The modified SCSR (SCSR-M) is available at https://doi.org/10.17605/OSF.IO/V6NTC.

Social Validity. Given the reliance on self-reported data, assessing social validity is crucial for evaluating the practical significance and acceptability of the PEERS®-YA social skills intervention among participants and key stakeholders. This study will use multiple methods for this assessment. First, as a post-intervention measure, I administered a modified version of the *Children's Intervention Rating Profile* (CIRP). The original CIRP was designed to assess the perceived fairness and expected effectiveness of educational interventions for children, using a 7-item Likert-type rating scale. For the original instrument, total scores are summed, with higher scores indicating greater levels of acceptability of the intervention. The CIRP has demonstrated high internal consistency (scores ranging from .75 to .89; Carter, 2009). For this study, I adapted the CIRP into a 10-item measure, updating its language and contextual relevance for university students (e.g., considering its effect on university life and its utility in navigating social situations). This modified version (CIRP-M) is available at https://doi.org/10.17605/OSF.IO/V6NTC.

Interviews. After the semester ended, participants willing to share their experiences were invited to semi-structured social validity interviews designed to explore their perceptions of the program's fairness, usefulness, and overall impact. Because time constraints precluded transcription, coding, and thematic triangulation of these interviews within the completed study, such qualitative analyses will be conducted and reported in a forthcoming manuscript. By later integrating emergent themes from these interviews with other data collected during this study, I anticipate achieving a more comprehensive, triangulated assessment of the intervention's acceptability and identifying areas for refinement to enhance its effectiveness.

Intervention. The independent variable in this study was the PEERS®-YA curriculum, used to assess its impact on the primary dependent variable: social engagement. Developed initially as PEERS® for Adolescents, the framework employed didactic methods to teach and practice essential social skills for establishing and maintaining interpersonal relationships. Although

PEERS®-YA has been less commonly applied in postsecondary settings, it has demonstrated significant positive outcomes in multiple studies across various contexts (Chien et al., 2021; Honan et al., 2023; Howorth et al., 2022; Howorth et al., 2023; Hua et al., 2022; Laugeson et al., 2015; McVey et al., 2016; McVey et al., 2017; Oh et al., 2021; Platos et al., 2023; Rose et al., 2021).

For this study, the intervention utilized a modified version of PEERS®-YA, focusing primarily on seven lessons from the curriculum. To tailor the intervention to participants' specific needs, lessons on dating and bullying were excluded, though students were offered the option to complete these sessions after the study concluded. Intervention sessions occurred weekly for ~45 minutes to align with participant schedules, and each individual received sessions of equal length to maintain consistency. The intervention spanned approximately seven weeks, with some lessons being combined into single sessions to accommodate the academic calendar.

By offering a combination of didactic lessons and role-play scenarios, participants learned and practiced a variety of skills, including but not limited to (a) developing and maintaining friendships, (b) conversational techniques such as entering and exiting conversations, (c) appropriate use of humor, (d) effective electronic communication, (e) hosting social events, and (f) managing disagreements (Laugeson, 2017).

Fidelity of Implementation. Treatment fidelity is crucial for ensuring that an intervention is implemented consistently and accurately as designed, thereby maximizing its potential effectiveness. Given the constraints of this study, modifications were made to the PEERS®-YA curriculum and its delivery to facilitate dissertation completion. The total number of sessions was reduced to seven, with some sessions combining lesson objectives, and individual session length was shortened from 90 minutes to 45 minutes to better align with course schedules and minimize participant time burdens. Due to limited resources, the study did not include the "social coaching" component that typically supports the intervention outside of direct session contexts. Despite these

modifications, the primary format of the intervention remained intact. Each session included *Homework Review*, *Didactic Lessons and Role-play Demonstrations*, and *Behavioral Rehearsal Exercises*. These adaptations were documented in an observational fidelity checklist to monitor the consistency of the intervention and aim to ensure participants received its core benefits; this checklist is available at https://doi.org/10.17605/OSF.IO/V6NTC.

Experimental Design and Procedures

This study employed a single-case experimental design (SCED) to test the effects of the PEERS®-YA curriculum on social engagement for autistic undergraduates. In addition, I measured social skills knowledge and an overall sense of belonging before and after the intervention. While these data were not experimentally tested in the SCED, they provided evidence of social validity and contextualized the changes observed in the time series data.

In this SCED, I used a multiple baseline design across participants to assess the intervention's efficacy and generalizability in enhancing social engagement. Multiple baseline designs across participants offer several advantages in research and applied settings by introducing the treatment at different times across individuals and settings, allowing for a temporal sequencing of repeated measurements (Hott et al., 2021; Pustejovsky et al., 2014). Given the small sample size and the focus on enhancing social understanding and skills, this design enabled me to collect data in natural settings during the baseline phase, monitor behaviors during the intervention phase, and conclude with observations during the maintenance phase. This systematic approach allowed for the observation of the intervention's effects across different participants and settings (Hott et al., 2021; Pustejovsky et al., 2014; Rizvi & Nock, 2008). Lastly, a multiple baseline design across participants was selected due to concerns that the dependent variable was potentially irreversible (i.e., an increased sense of belonging, once experienced, would not quickly regress), which eliminated the applicability of a withdrawal (ABAB) design.

Additionally, the multiple baseline design generally facilitates staggered introductions of the intervention. This approach enables comparisons of behaviors across participants, reduces the likelihood of baseline interdependence, and increases confidence that observed effects are due to the intervention rather than extraneous factors (Bulté & Onghena, 2009; Doggett et al., 2013; Sanson-Fisher et al., 2014; Willson et al., 2016). As this study was part of a dissertation, it was conducted over a single semester (14 to 16 weeks). Baseline data collection periods across participant groups were systematically staggered at 3, 5, and 7 weeks to ensure adequate data collection at each phase.

Prior To Data Collection

After IRB approval, participants provided informed consent using a standard, IRB-approved signature form. During this process, students were asked to consent to audio and video recordings of intervention sessions, provided it was deemed necessary for treatment fidelity. The consent forms included an option for participants to opt out of recordings and a provision allowing them to withdraw from the study at any time. I scanned and stored the consent forms on an encrypted, university-provided drive, with paper copies securely stored in a locked cabinet within a secured office on campus.

Although all students in the on-campus support program had the opportunity to receive the intervention, participation in this study was entirely voluntary and was not required to maintain their eligibility for the program. I also ensured that participants understood they could withdraw from the study at any time without any negative consequences or loss of eligibility for program participation in future academic years.

Training for Staff Roles. Due to staffing limitations, this study was conducted without the typically included social coaching group (e.g., parents, peers), and data collection focused specifically on the young adult participants. As the lead researcher delivering the intervention, I had completed both the International Board of Credentialing and Continuing Education Standards (IBCCES)

Autism and Neurodiversity Training and Certification Course and the PEERS®-YA Certification Course. To help ensure treatment fidelity, a graduate researcher attended and documented student attendance and adhered to a treatment fidelity checklist. Had the graduate researcher been absent, I intended to record the sessions myself to ensure treatment fidelity could still be reported; however, we were both present for all sessions with both participants.

Training and Assessment for Participants. Given this study's reliance on technology for collecting self-reported data, it was essential to provide participants with the necessary tools and training to receive prompts and submit responses effectively. During the first week of the semester, I met with participants to create their *m-Path* accounts and demonstrated the process for receiving and responding to daily prompts. This initial meeting also involved administering pre-baseline assessments (i.e., SCSR-M, TYASSK) to help establish a data collection routine and gather pre-intervention data. During the first week of the baseline phase, I reinforced participants' understanding and addressed any questions or concerns they had regarding the data collection process.

Data Collection

ESM is a momentary assessment technique that offers significant methodological advantages to educational researchers by enabling the assessment of individuals in their natural environments across multiple time points. Other terms for ESM include ecological momentary assessment, ambulatory assessment, event sampling, daily diary, intensive longitudinal assessment, and real-time data capture studies (Bentley et al., 2021; Mölsä et al., 2022; Verhagen et al., 2016). ESM provides flexible, personalized questionnaire content delivered through time-, signal-, or event-contingent schedules. This approach achieves high ecological validity by studying moment-to-moment variations in affect influenced by daily experiences in real-life contexts (Atz, 2013; Bentley et al., 2019; Fullagar & Kelloway, 2009; Mazereel et al., 2021; Mölsä et al., 2022; Verhagen et al., 2016).

These schedules can occur at predetermined time points, random or semi-random intervals, or be user-initiated after specific events, enhancing accessibility through delivery to personal electronic devices such as smartphones or smartwatches (Bentley et al., 2019; Mölsä et al., 2022; Verhagen et al., 2016). ESM captures dynamic fluctuations in affect, cognition, behavior, and social contexts through real-time assessments of participants' "at the moment" responses. This immediacy minimizes recall bias and avoids data aggregation by repeatedly assessing specific points in time, thereby offering more ecologically valid self-report measures (Atz, 2013; Bentley et al., 2019; Verhagen et al., 2016; von Klipstein et al., 2023).

Repeated measurement is a fundamental characteristic of ESM, which facilitates the collection of robust data sets across multiple contexts while minimizing errors typically associated with time series data collection, thus enhancing the validity, reliability, and transparency of the collected data (Atz, 2013; Bentley et al., 2019; Fullagar & Kelloway, 2009; Verhagen et al., 2016). The intensive longitudinal design inherent in ESM enables researchers to track changes in targeted variables throughout a study by collecting repeated measures across all study phases (e.g., Baseline, Intervention, Maintenance), thereby providing a more comprehensive understanding of the targeted variables (Atz, 2013; Fullagar & Kelloway, 2009; Verhagen et al., 2016).

This real-time data collection aligns seamlessly with the principles of SCED, which require (a) systematic and repeated measurement across all phases to bolster internal validity, (b) an adequate number of observations to facilitate appropriate statistical analyses, particularly given the small sample sizes typical in SCED, (c) sufficient data points to evaluate the effectiveness of the intervention in producing behavior change, and (d) the flexibility to make data-informed decisions regarding the independent variables (Bentley et al., 2019). Furthermore, ESM's capacity for real-time monitoring outside of clinical or laboratory settings addresses a key limitation of traditional SCED studies by enhancing the ecological validity of the findings. This naturalistic observation provides

deeper insight into behavior and strengthens the generalizability of the results to real-world contexts (Bentley et al., 2019).

Given the significant personal responsibility and time commitment required in ESM studies, participants may experience an increased burden, particularly within SCEDs, which often involve a smaller pool of participants committing to extensive data collection (Fritz et al., 2024; Verhagen et al., 2016). To enhance participant compliance and minimize the potential for missing data, researchers should design repeated measures that can be completed within one to three minutes, reducing the time investment for participants (Fritz et al., 2024; Verhagen et al., 2016).

Another critical consideration in ESM studies is the delivery of prompts, with most studies relying on fixed, random, semi-random, or event-contingent schemes. While event-contingent schemes are beneficial for capturing infrequent events, such as social interactions, they impose a continuous monitoring burden on participants and may not effectively reduce recall bias or retrospection (Fritz et al., 2024; Verhagen et al., 2016). In contrast, while often correlated with higher participant engagement, fixed schedules introduce predictability, allowing participants to anticipate prompts and potentially alter their behavior in response. This predictability can compromise the representativeness of the sample by inducing reactivity (Fritz et al., 2024; Verhagen et al., 2016; von Klipstein et al., 2023).

In a fully randomized schedule, extended periods without prompts can disrupt participants' routines, leading to skewed data as they may alter their daily activities to ensure they receive the prompt. Semi-random or stratified random schedules are recommended to balance participant burden with data quality, as they provide a more manageable schedule while ensuring sufficient and diverse data collection (Fritz et al., 2024; Verhagen et al., 2016).

For this study, I implemented daily assessments with a target completion time of ≤3 minutes to minimize participant burden. The timing of these assessments was randomly determined using the

RAND function in Microsoft Excel, with prompts scheduled to occur at two of three predetermined times throughout the day (10:00 AM, 2:00 PM, and 6:00 PM). Each student received prompts at distinct times to introduce variability across groups. Additionally, to ensure the capture of participants' experiences and emotions in real-time, a reminder prompt was issued 60 minutes after the initial prompt, and the completion window closed 120 minutes following the original prompt. This approach was designed to enhance the ecological validity of the data by reducing recall bias and increasing self-report accuracy, while also providing participants enough time to complete the survey if a prompt occurred when they were otherwise occupied (e.g., class, lab, work, leisure). Furthermore, I set prompts to fire randomly within the data collection window to prevent participants from developing a predictable schedule or circumventing notifications.

Baseline Phase

Baseline data was collected through daily prompts sent to participants' electronic devices (e.g., smartphones, smartwatches). Participants received two prompts throughout the day asking the predetermined survey questions about their sense of belonging (See Appendix B: Figure 1). Each baseline data point represented the weekly mean or median of these daily prompts, calculated from individual responses.

Intervention Phase

Due to time constraints, the intervention commenced after three weeks of baseline data collection, with the introduction of the PEERS®-YA curriculum. Each participant began the intervention at different intervals, allowing for systematic observation of the intervention's effects over time. The first participant started the intervention immediately after their initial three-week baseline period, the second began two weeks later, and the third commenced two weeks after the second. This staggered introduction facilitated continuous monitoring and comparison of social

engagement across individuals, providing insights into the intervention's effectiveness throughout the study.

Maintenance Phase

To accommodate participant requests to complete all study activities before final, the planned maintenance phase was modified. For each participant, maintenance data collection commenced immediately after their individual intervention sessions concluded and was completed entirely before exams began, rather than occurring as a distinct, extended post-semester period. During this adapted maintenance phase, participants provided self-reported data via the regular bi-daily probes that focused on measures of social engagement, sense of belonging, and positivity. The study aimed to evaluate, through these data, how effectively participants maintained and applied the social skills learned during the intervention in their daily lives within this immediate post-intervention timeframe. This assessment was intended to gauge the short-term efficacy of the modified PEERS®-YA program and its contribution to the students' adjustment to university life. Concurrently with this maintenance data collection, participants completed post-intervention assessments of the dependent measures, and social validity measures were gathered to ascertain their perspectives on the intervention's acceptability and perceived effectiveness.

Treatment Fidelity

Treatment fidelity was rigorously upheld by strictly adhering to the PEERS®-YA curriculum framework, with adjustments made only when deemed essential. I, as the sole facilitator credentialed in both IBCCES and PEERS®-YA, led each intervention session, ensuring a high level of expertise and consistency in delivery. A graduate researcher, serving as an independent observer, regularly conducted checks (See Appendix B: Figure 2) to verify that each session adhered to the prescribed curriculum, further safeguarding fidelity. Any deviations from the curriculum were meticulously

documented and reported; unintentional variations, had they occurred, would have been promptly addressed through additional training or procedural adjustments.

Rigorous monitoring of treatment fidelity is essential to ensure that any observed changes in the dependent measures can be confidently attributed to the intervention, thereby enhancing the validity and reliability of the study's findings. As this study was part of a dissertation, specific modifications to the intervention (e.g., shortening session lengths to accommodate participant schedules and tolerance, reducing the number of sessions from the recommendation, omitting the social coaching component) were implemented, as detailed previously. These adjustments were thoroughly documented and are discussed in the limitations section of the dissertation.

Preliminary Data Analyses

The time-series data was evaluated in four ways (a) visual analysis, (b) structured visual analysis, (c) non-overlap quantitative index, and (d) within-case quantitative index.

Visual Analysis. Visual analysis was the initial method used for interpreting data in this study, focusing on changes in student social engagement's effect on an individual's sense of belonging and overall feelings of positivity. Employing a multiple baseline design across participants allowed for the intervention to be introduced at staggered intervals. This design facilitated a detailed observation and analysis of the intervention's impact by examining key features within and between phases, (e.g., immediacy of effect, level, trend, variability). During each phase of the study (i.e., *Baseline*, *Intervention*, and *Maintenance*), each participant's responses were systematically evaluated. This analytical approach is grounded in the six core elements of visual analysis as outlined by Ledford et al. (2017), which provide a comprehensive framework for interpreting behavioral changes. These elements are the level of the data within phases, the trend or direction of behavior over time, and the variability of data points, as well as the immediacy of effect, the degree of overlap between data in adjacent phases, and the consistency of data patterns when comparing across phases.

The ***level*** refers to the amount or degree of the behavior that occurs within a specific phase, typically measured by the overall value (e.g., mean, median) of the data points, reflecting the consistency or stability of performance during that period. A stable level is considered important for making reliable predictions about future behavior, assuming conditions remain unchanged (Ledford et al., 2017).

The ***trend*** (or slope) indicates the direction of movement in the data over time within a phase, such as accelerating (increasing), decelerating (decreasing), or showing no clear direction (zero-celerating). A trend is considered *therapeutic* if it aligns with the intended outcome of the intervention or *contra-therapeutic* if it moves against the desired effect. Analyzing the trend within a phase is crucial for determining appropriate times to introduce or change conditions (Ledford et al., 2017).

Variability refers to the extent to which data points differ from each other within a phase, serving as a measure of data stability. Data are considered stable if approximately 80% of the values fall within ±25% of the median (the "stability envelope"). High variability can pose challenges in interpreting data, potentially obscuring whether changes in level reflect true effects or natural fluctuations. If variability was high during a baseline phase in this study, continuation of the phase would have been considered to achieve stability, unless a strong intervention effect that could overcome this variability was anticipated (Ledford et al., 2017).

Between phases, visual analysis examined the immediacy of effect, non-overlap of data (or its inverse, overlap), consistency of data patterns, and changes in level, trend, and variability.
The ***immediacy of effect*** refers to how quickly the behavior changes after introducing a new condition, often assessed by the change in level between the last few data points of one phase and the first few of the next. While immediate, clear shifts are preferable, delayed or gradual shifts

require careful monitoring to determine if the pattern is expected for the intervention or influenced by other factors (Ledford et al., 2017).

Overlap of data assesses the extent to which data from an intervention phase differ from data in an adjacent baseline phase. Conversely, significant overlap (where data points from one phase are at the same level as those in an adjacent phase) can reduce confidence in an intervention's impact. The greater the non-overlap, the stronger the evidence of the intervention's effect (Ledford et al., 2017).

Consistency of data patterns involves tracking similar data patterns (i.e., replication of changes in level, trend, and/or variability) within like conditions, such as across baseline conditions for different participants in a multiple baseline design. Consistent changes, even if modest, build confidence in the intervention effect, whereas inconsistent changes compromise such confidence (Ledford et al., 2017).

Changes in level, trend, and variability between phases were compared to determine the intervention's impact. This involved assessing how much the behavior altered (changes in level), whether the direction and rate of change shifted (changes in trend), and whether the consistency of performance improved or worsened (changes in variability). For instance, a shift from high variability in baseline to more consistent data during intervention could indicate a positive intervention effect (Ledford et al., 2017; Kratochwill et al., 2010).

Structured Visual Analysis. To corroborate the visual analysis, I used a structured approach, specifically the conservative dual-criterion (CDC) method (Fisher et al., 2003). The CDC has been shown to correlate strongly with expert visual analysis and provides an objective means of assessing treatment effects (Swoboda et al., 2010). Application of the CDC in this study involved creating two criterion lines based on baseline data: **(a)** a horizontal line set at 0.25 standard deviations (SD) of the baseline data beyond the mean of the baseline phase, in a therapeutic

direction; and **(b)** a line parallel to the baseline trend line, but displaced from the projected baseline trend by 0.25 SD of the baseline data, also in a therapeutic direction (Fisher et al., 2003). Subsequently, the number of intervention data points falling beyond both criterion lines was counted, considered in relation to the total number of intervention data points, and compared to a binomial distribution to determine statistical significance. To compute the CDC, I used Microsoft Excel, following a tutorial and macros package provided by Swoboda et al. (2010).

Non-Overlap of All Pairs. Non-overlap of All Pairs (NAP) is an index of overlapping data across phases in SCEDs and has a strong presence and history as an indicator of performance differences typically assessed through visual analysis in single-case studies. NAP measures the degree of data overlap by comparing every data point in the baseline phase with every data point in the intervention phase. This comprehensive approach provides a robust and sensitive metric for detecting changes attributable to the intervention. NAP has advantages over other non-overlap indices, such as Percentage of Nonoverlapping Data (PND) and Percentage of Intervention Data Points Exceeding the Median of Baseline (PEM), because it offers greater discriminative power and aligns more closely with visual analysis judgments. Additionally, NAP has been shown to correlate strongly with established effect size measures like *Pearson's* R^2, enhancing its validity. By utilizing NAP, this study aimed to achieve a more precise and reliable evaluation of the intervention's impact on participants' social engagement (Parker & Vannest, 2009). To calculate NAP, I used the Single-Case Effect Size Calculator (Version 0.7.3; Pustejovsky et al., 2024), available at https://jepusto.shinyapps.io/SCD-effect-sizes/.

Within-Case Quantitative Index. The log response ratio (LRR) was employed in this study to quantify the effect size between the *Baseline* and *Intervention* phases for each participant in the SCED. The LRR measures proportional change by calculating the natural logarithm of the response ratio, which is the ratio of the mean level of the outcome during the Intervention phase to the mean

level during the baseline phase (Pustejovsky, 2018): $LRR=\ln\left(\frac{M_I}{M_B}\right)$. This effect size is particularly appropriate for outcomes measured on a ratio scale where zero corresponds to the true absence of the outcome (Pustejovsky, 2018). There are two versions of LRR: LRR_i, associated with positive values suggesting a therapeutic trend, and LRR_d, associated with negative values indicating a contra-therapeutic trend. For outcomes measured as frequency counts or rates, LRR_i and LRR_d have the same magnitude but opposite signs; however, for outcomes measured as percentages or proportions, they differ in both sign and magnitude (Pustejovsky, 2018). By employing the LRR, this study aimed to capture multiplicative effects and proportional changes in participants' social engagement's affect sense of belonging and overall feelings of positivity. This metric is particularly suitable for data exhibiting skewness or proportional changes, common in social and behavioral research. To calculate LRR, I used the Single-Case Effect Size Calculator (Version 0.7.3; Pustejovsky et al., 2024), available at https://jepusto.shinyapps.io/SCD-effect-sizes/. Given this intervention's intent to establish a therapeutic trend for both belonging and positivity, I used LRR_i when calculating results.

Pre- and Post-Assessment Data. Raw change scores were computed for each participant by subtracting their pre-intervention total score from their post-intervention total score on both the SCSR-M and the TYASSK. Individual difference scores are presented in the **Results** section, accompanied by a concise narrative summarizing how the intervention influenced each student's sense of belonging and social skills knowledge.

Study Deviations

Initially, three undergraduates meeting inclusion criteria were recruited and provided informed consent to participate in this study; however, one participant suffered an injury unrelated to the study procedures early in their baseline phase (Week 2), resulting in their withdrawal from the study. Data collection for this individual was discontinued, and their data were not included in the final analyses reported herein. Consequently, the study proceeded with the two remaining

participants (i.e., Chase, Alyx), and the multiple baseline across participants design was implemented across these two individuals instead of the initially planned three replications.

During the intervention phase, two notable disruptions to the schedule occurred. First, the university closed for a week-long holiday (i.e., Spring Break; Week 7), during which no intervention sessions were held. Second, I fell ill during Week 9, which necessitated a one-week postponement of intervention sessions, as I was the sole individual certified to deliver the PEERS®-YA curriculum. This schedule adjustment ultimately resulted in one less week than originally planned for the collection of maintenance data.

Results

Results are reported by participant, looking at traditional graphs, assessment data, and any potential comments left by individuals.

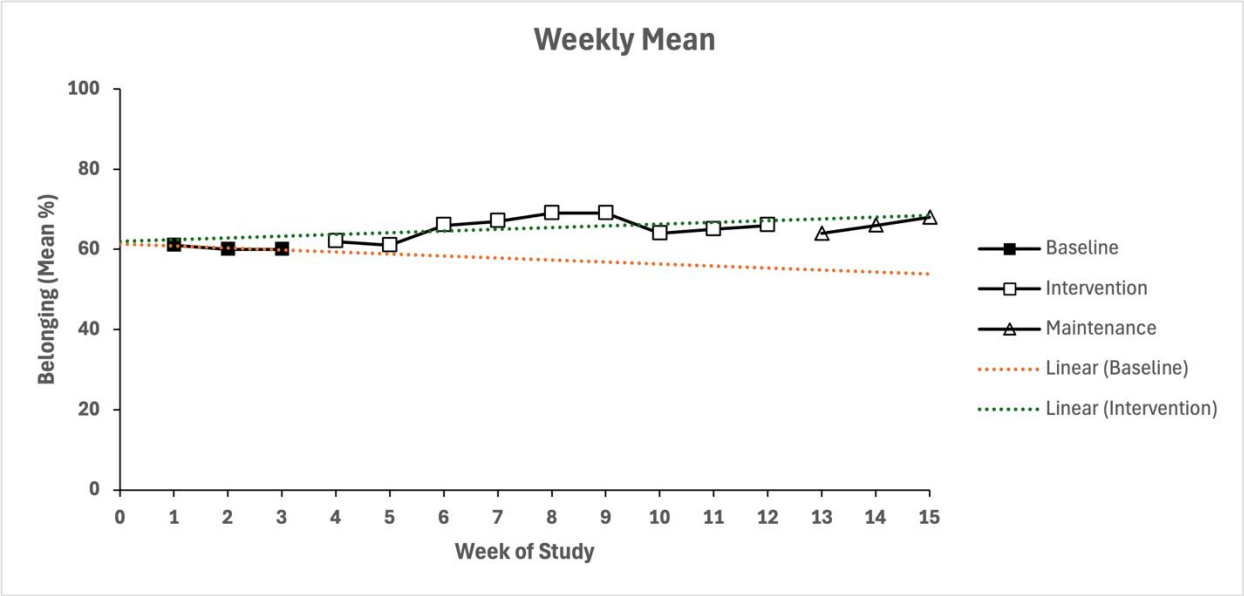
Chase

Belonging

Preliminary Visual Analysis. Visual analysis of Chase's weekly mean belonging scores, employing the six criteria from Ledford et al. (2017; See Figure 3), demonstrated a clear and replicable effect of the supports. Initially, during the baseline phase, his scores exhibited minimal variability alongside a slight downward trend. With the introduction of supports in Week 4, there was an immediate elevation in mean belonging, which then ascended consistently throughout the intervention phase, establishing an upward trajectory well above baseline levels while maintaining stable variability. A transient decrease in Week 10, possibly linked to academic pressures like midterms or other workload demands, underscored the potential benefits of exploring such anomalies through qualitative methods, (e.g., participant interview, semester schedule). Upon withdrawal of supports in Week 13, scores showed an initial decline, akin to the baseline pattern, but subsequently rebounded and continued to improve during the maintenance phase, thereby

preserving most of the gains achieved during the intervention. Overlap between phases was negligible, with only a single intervention data point falling within the baseline's range. Collectively, the consistent, phase-contingent changes across level, trend, variability, immediacy, overlap, and replication provided robust visual evidence that the intervention reliably enhanced Chase's sense of belonging.

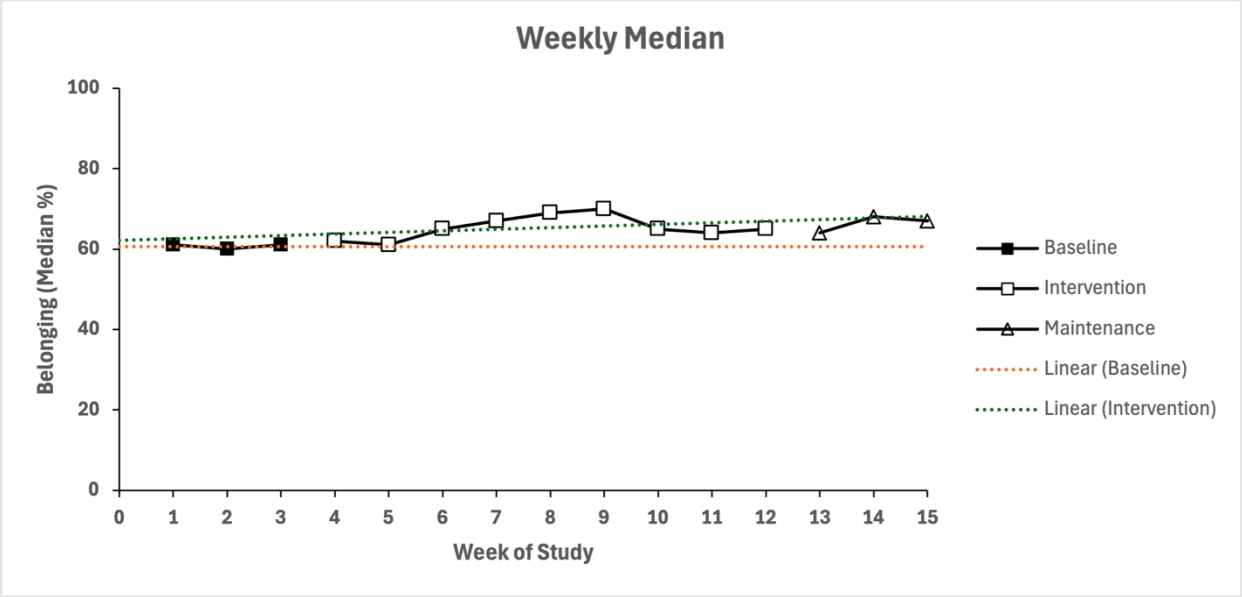
Figure 3. Chase's Belonging (Weekly Mean)



Further examination of Chase's weekly median belonging scores (See Figure 4), utilizing the same six visual-analysis criteria (Ledford et al., 2017), reinforced the finding of an apparent intervention effect, though the timing of this effect presented slightly differently compared to the mean. Baseline median scores were characterized by tight clustering, indicating negligible variability and a flat trend. Following the introduction of supports in Week 4, Chase's median sense of belonging began to rise, climbing steadily to its peak in Week 9. Similar to the mean analysis, a minor dip was observed in Week 10, potentially attributable to temporary academic stressors or other confounding variables, which highlighted an opportunity for deeper understanding via qualitative investigation (e.g., follow-up interviews, academic schedule analysis). After supports were withdrawn in Week 13, the median scores initially fell but then showed a recovery throughout the maintenance

phase, retaining much of the progress made during intervention. Variability remained consistently low across all phases, and overlap was minimal, with only one intervention data point aligning with the baseline range. The observed patterns of change in level, trend, variability, immediacy, and the degree of overlap and replication across phases strongly echoed the mean-based findings. However, the mean scores captured the intervention's initial impact more rapidly, suggesting that while the mean may be more sensitive to immediate shifts, the median offered a robust confirmation of the central tendency of change, potentially less swayed by isolated deviant data points.

Figure 4. Chase's Belonging (Weekly Median)



When plotted across all daily prompts (n = 178; See Appendix B: Figure 3), Chase's raw belonging scores fluctuated around the same overall level and trend observed in the weekly summaries but revealed finer-grained variability. The full-series graph showed a modestly upward drift over time, punctuated by abrupt shifts at the onset and withdrawal of supports that mirrored the week-by-week mean and median inflections. Although the weekly mean analysis had detected the intervention effect one prompt sooner and the weekly median had provided a robust, outlier-resistant confirmation, this daily plot distinctively highlighted intermittent dips and spikes that were averaged out in the aggregated weekly summaries. Regardless, the same phase-linked increases,

reversals, and replications appeared in the daily data, confirming that aggregating to weekly means or medians appropriately captured the intervention's core effect while smoothing day-to-day variability.

Structured Visual Analysis. Both the weekly mean and the weekly median summaries produced identical conservative dual-criterion (CDC) indices (See Table 6), with seven of nine treatment-phase data points falling above both the shifted mean and shifted trend lines. However, eight "successes" were required to exceed the binomial threshold ($p < .05$), meaning neither the mean-based nor the median-based analysis reached the criterion for a statistically reliable effect according to this method. Notably, the fact that the mean and median yielded the same CDC outcome suggested that extreme values had minimal influence on this decision index in this case. Extending the CDC procedure to the complete daily dataset was impractical using the available spreadsheet macros, as it would have required first adjusting the embedded formulas to accommodate the larger volume of data. Future work will involve revising these macros or porting the computation to a statistical package (e.g., *R*) to enable such full-data assessments.

Table 6. Chase's Conservative Dual-Criterion (Belonging)

Weekly	# of Treatment Points	# Above Both Lines	# Needed Above Both Lines
- Mean	9	7	8
- Median	9	7	8

Non-Overlap of All Pairs and Within-Case Quantitative Index. For Chase, the transition from baseline to intervention (See Table 7) yielded a Non-overlap of All Pairs (NAP) of .89 (95% CI [.81, .94]), indicating that 89% of his intervention-phase belonging scores exceeded those observed during baseline. The corresponding Log Response Ratio (LRR) was 0.08 (95% CI [0.07, 0.10]), reflecting a proportional increase of approximately 9% (95% CI [7%, 10%]) in his mean belonging score once supports were introduced. When supports were withdrawn and the maintenance phase began, the NAP comparing the intervention to maintenance phases fell to .53 (95% CI [.42, .64]), signifying only slight non-overlap. Concurrently, the LRR comparing these

phases dropped to 0.01 (95% CI [–0.01, 0.02]), equivalent to a negligible 1% change (95% CI [–1%, 2%]). Notably, although Chase's sense of belonging dipped when supports were withdrawn, it did not return to baseline levels; his maintenance-phase scores remained above pre-intervention values, suggesting that he retained a portion of the gains even without ongoing active supports.

Table 7. Chase's Belonging (NAP; LRR)

Domain	Phase Shift	NAP [95% CI]	LRR [95% CI]	% Change [95% CI]
- Belonging	BL>>>INT	0.89 [0.81, 0.94]	0.08 [0.07, 0.10]	9 [7, 10]
- Belonging	INT>>>MN	0.53 [0.42, 0.64]	0.01 [-0.01, 0.02]	1 [-1, 2]

Confounding Variable: Social Partner. Table 8 displays Chase's self-reported belonging scores disaggregated by the presence or absence of a social partner during each bi-daily prompt. Across all phases, the presence of a social partner was associated with slightly higher belonging ratings, though the magnitude of this effect varied. During the intervention phase, belonging scores were higher when a social partner was present (M = 66.7) than when one was not (M = 64.5), a difference of 2.2 points. A similar pattern was observed in the maintenance phase, though with a smaller difference (M = 66.1 with a partner vs. M = 65.9 without). Interestingly, during the baseline phase, scores were slightly higher when no partner was present (M = 61.2) than when one was (M = 60.0), suggesting that early patterns of social engagement may have differed from those observed later. Overall, while the presence of a social partner appeared to modestly elevate Chase's belonging ratings during the intervention and maintenance phases, the size of these differences was small and seemed unlikely to account for the full pattern of change observed in his overall belonging scores.

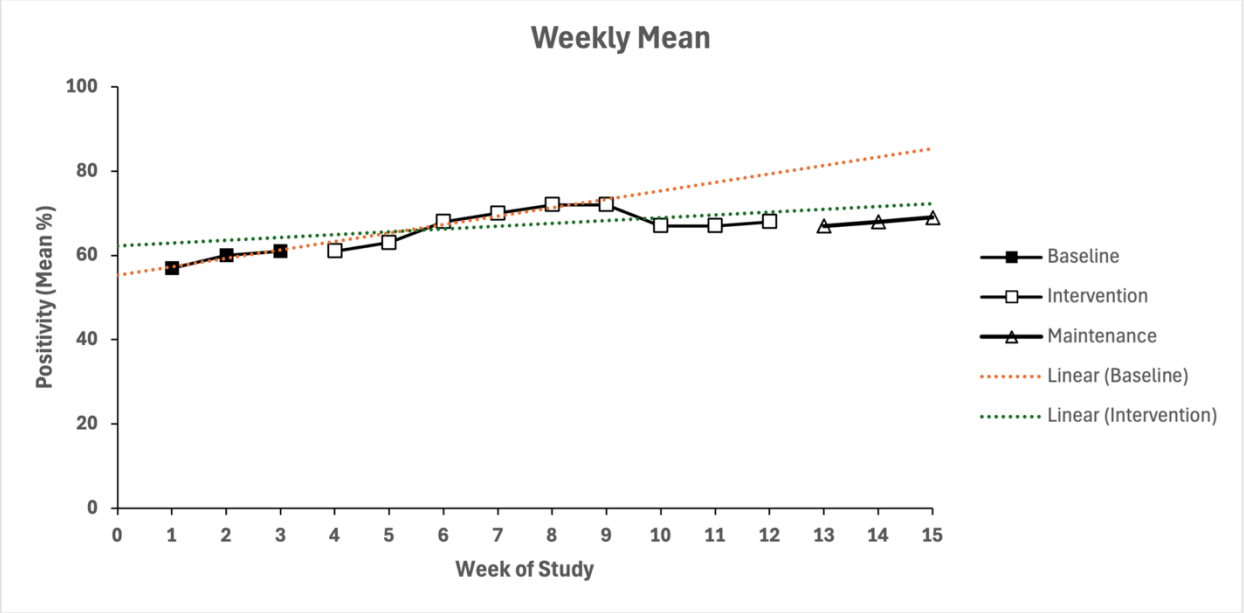
Table 8. Chase's Social Partner Report (Belonging)

Condition	Social Partner	N	Mean	Median	SD	Min	Max
- Baseline	Y	12	60.0	60.0	2.14	58	65
- Baseline	N	26	61.2	61.0	2.33	56	64
- Intervention	Y	53	66.7	66.0	3.27	59	76
- Intervention	N	52	64.5	64.5	3.51	58	70
- Maintenance	Y	10	66.1	66.0	2.28	62	69
- Maintenance	N	25	65.9	65.0	2.93	61	72

Positivity

Preliminary Visual Analysis. Analysis of Chase's weekly mean positivity scores (See Figure 5), guided by the six visual-analysis criteria (Ledford et al., 2017), revealed a pronounced and replicable effect of the supports. During the baseline period, mean positivity showed a modest, gradual increase characterized by minimal variability and a slight upward trend. Upon the introduction of supports in Week 4, the scores, after initially holding steady, began to climb markedly, achieving a notable elevation compared to baseline levels, all while maintaining a stable variability envelope. A distinct drop in Week 10, possibly reflecting transient academic pressures or other contextual influences, pointed to the utility of further qualitative exploration, perhaps through participant interviews or an examination of the semester's schedule. Following the withdrawal of supports in Week 13, mean positivity decreased, briefly mirroring the baseline pattern. which underscored the immediacy of effect when supports were removed, but then recovered strongly in Weeks 14 to 15, thus preserving many of the gains from the intervention. With minimal overlap between phases (i.e., only one intervention data point fell within the baseline range), the consistent reversals in level and trend observed at each phase transition, combined with the overall patterns of variability, immediacy, and replication, offered compelling visual evidence that the supports reliably enhanced Chase's sense of positivity.

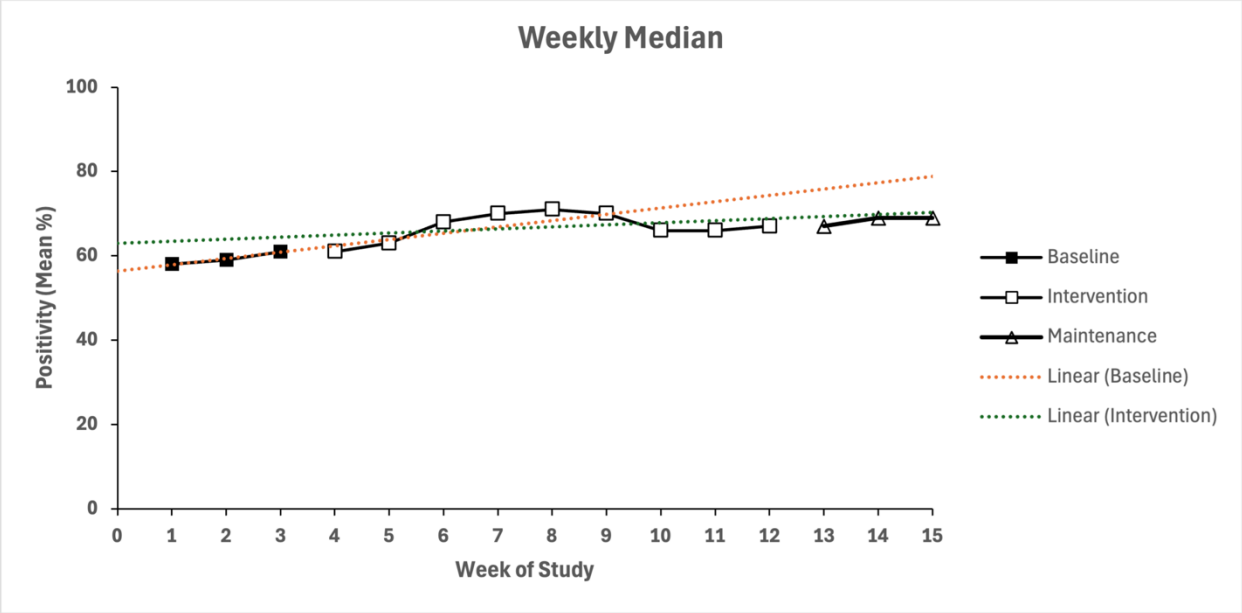
Figure 5. Chase's Positivity (Weekly Mean)



The pattern of Chase's weekly median positivity scores (See Figure 6), when visually inspected using Ledford et al.'s (2017) six criteria, also indicated a clear intervention effect, though its emergence differed subtly from the mean-based analysis. Baseline median positivity was notable for its tight clustering of data points, which translated to negligible variability around a slight upward trend. Once supports were introduced in Week 4, the medians rose without delay and ascended steadily throughout the intervention phase, peaking in Weeks 8 to 9. The minor downturn in Week 10 was conjectured to be a result of temporary academic demands or other situational confounds, a hypothesis warranting deeper investigation through qualitative means (such as discussions with the participant or a review of contemporaneous university events). After the supports were discontinued, median scores experienced an initial dip before rebounding during the maintenance phase, thereby retaining most of the benefits accrued during the intervention. Data variability remained low across all phases, and only a single intervention data point overlapped with the baseline range. These phase-contingent shifts in level, trend, and variability, along with considerations of immediacy, overlap, and replication, largely corroborated the findings from the

mean-based graph. It was observed, however, that the mean registered the initial effect more rapidly; this suggested that while the mean could be more sensitive to immediate changes, the median provided a robust confirmation of the central tendency, appearing less influenced by outlier data points.

Figure 6. Chase's Positivity (Weekly Median)



When plotted at the level of all received daily prompts (n = 178; See Appendix B: Figure 4), Chase's positivity ratings revealed a gentle upward drift over time, punctuated by subtle upticks immediately following the introduction of the intervention and slight downturns upon its withdrawal; these patterns mirrored those observed in the weekly summaries. The raw daily series was noticeably more volatile than the weekly aggregates, exhibiting momentary spikes and troughs that the weekly mean (which had flagged the effect earlier) and the weekly median (which had offered a stable, outlier-resistant confirmation) smoothed over. Yet despite this fine-grained "noise," the characteristic phase-linked boosts, reversals, and replications in level and trend remained evident in the daily data. This demonstrated that weekly mean or median aggregation faithfully captured the core intervention signal while filtering out day-to-day variability.

While the intervention did produce a therapeutic effect, it is essential to note that Chase's positivity had already shown a therapeutic trend, indicating a relatively high starting point.

Structured Visual Analysis. Application of the CDC method to Chase's weekly positivity summaries (See Table 9) did not yield findings indicative of a statistically reliable intervention effect for either the mean or median data. Specifically, for the weekly mean positivity scores, none of the nine treatment-phase weeks surpassed both the therapeutically shifted mean and trend lines. The weekly median positivity data showed a slightly different outcome, with one of the nine median points meeting this dual criterion. However, given that eight such "successes" were needed to meet the binomial threshold for statistical significance ($p < .05$), neither aggregation method (mean nor median) satisfied the requirements for the CDC. This minor difference in the count of successes between the mean (zero) and median (one) data further suggested that isolated high or low scores within the weekly data did not substantially impact the overall conclusion from this specific decision-making index. Applying the CDC to the full daily positivity dataset was not feasible with the existing spreadsheet macros due to the extensive formula modifications that would have been necessary for the larger data volume. Future work will involve exploring alternative computational approaches, such as custom scripting or the use of specialized functions within statistical software (e.g., *R*), to facilitate a CDC assessment of the complete daily data series.

Table 9. Chase's Conservative Dual-Criterion (Positivity)

Weekly	# of Treatment Points	# Above Both Lines	# Needed Above Both Lines
- Mean	9	0	8
- Median	9	1	8

Non-Overlap of All Pairs and Within-Case Quantitative Index. For Chase's positivity scores, the transition from baseline to intervention (See Table 10) yielded a NAP of .91 (95% CI [.84, .95]), a value indicating very high non-overlap between these phases. Concurrently, the LRR was 0.13 (95% CI [0.11, 0.16]), which corresponded to an approximate 14% proportional increase

(95% CI [11%, 18%]) in his mean positivity score once supports commenced. When examining the subsequent transition from intervention to the maintenance phase, a different pattern emerged: the NAP comparing these latter phases decreased substantially to .47 (95% CI [.37, .58]), while the LRR was effectively zero at 0.00 (95% CI [-0.02, 0.01]). This LRR value translated to essentially no sustained proportional change (0%; 95% CI [-2%, 1%]) from the intervention into the maintenance period. It is noteworthy, however, that while Chase's positivity dipped when supports were withdrawn, his scores did not revert to initial baseline levels; maintenance-phase positivity scores consistently remained above pre-intervention values, suggesting he preserved some of the benefits despite the absence of active supports.

Table 10. Chase's Positivity (NAP; LRR)

Domain	Phase Shift	NAP [95% CI]	LRR [95% CI]	% Change [95% CI]
- Positivity	BL>>>INT	0.91 [0.84, 0.95]	0.13 [0.11, 0.16]	14 [11, 18]
- Positivity	INT>>>MN	0.47 [0.37, 0.58]	0.00 [-0.02, 0.01]	0 [-2, 1]

Confounding Variable: Social Partner. Table 11 presents Chase's positivity ratings disaggregated according to the presence or absence of a social partner during each bi-daily prompt. In the baseline phase, whether a social partner was present had virtually no bearing on positivity ratings (M = 59.3 with a partner vs. M = 59.1 without); medians and variability were also nearly identical between these conditions. A slight shift occurred during the intervention phase, where positivity was marginally higher when Chase was with a social partner (M = 68.6) compared to when he was alone (M = 66.9), a difference of 1.7 points. However, this subtle advantage did not carry over into the maintenance phase. During maintenance, mean scores were identical (M = 67.8) irrespective of a social partner's presence, and other indicators such as median, variability, and range showed negligible differences. These findings suggested that while being with a social partner might have provided a minor boost to Chase's positivity during the active intervention period, this effect was minimal and did not endure once the primary supports were withdrawn. Therefore, although

social partner presence could be considered a minor situational factor, it did not appear to have meaningfully confounded the primary intervention effects observed on Chase's overall positivity.

Table 11. Chase's Social Partner Report (Positivity)

Condition	Social Partner	N	Mean	Median	SD	Min	Max
- Baseline	Y	12	59.3	60.0	6.07	50	69
- Baseline	N	26	59.1	60.0	4.19	50	68
- Intervention	Y	53	68.6	68.0	4.22	57	77
- Intervention	N	52	66.9	68.0	3.92	58	76
- Maintenance	Y	10	67.8	68.0	2.10	64	72
- Maintenance	N	25	67.8	68.0	2.10	63	73

Measurement

Social Connectedness Scale—Revised (Modified). Table 12 presents Chase's pre- and post-intervention ratings on the SCSR-M, which indicated an overall improvement in his reported sense of connection and belonging on campus. At baseline, Chase's mean score across the ten items was 60.0; at post-test, this score rose to 77.1, representing a mean gain of 17.1 points. Scores on nearly all items increased. The largest gains were observed on items such as "I find it easy to connect with other students and professors" (from 40 to 73, +33 points) and "I see other students and professors as approachable and friendly" (from 63 to 90, +27 points). These changes suggested that the supports most strongly enhanced his perceived ease of initiating new interactions and his perception of the campus community's welcoming nature. Substantial improvements also occurred on items related to general campus connectedness (+20 points) and comfort in interacting with unfamiliar peers (+35 points). One exception was a slight decrease on the item "I feel like I belong in social and academic settings" (from 80 to 76, -4 points). This finding indicated that while Chase's confidence in initiating connections appeared to increase, his underlying sense of belonging within mixed social-academic contexts may warrant further exploration. Overall, the pattern of change suggested that the intervention notably improved Chase's social engagement and perceived

accessibility of peers and faculty, even as specific aspects of academic belonging emerged as an area for potential ongoing support.

Table 12. Chase's Pre-Post (Revised Social Connectedness Scale–Modified)

	Question	Pre	Post
1.	I feel connected to other students on campus.	66	***80***
2.	I actively participate in campus groups or activities.	70	***77***
3.	I feel like I belong in social and academic settings at the university.	***80***	76
4.	I find it easy to connect with other students and professors.	40	***73***
5.	I feel understood and accepted by my peers and professors.	71	***84***
6.	I see other students and professors as approachable and friendly.	63	***90***
7.	I feel comfortable interacting with other students and professors I don't know well.	35	***70***
8.	I feel connected to the broader university community.	57	***77***
9.	The friends I have made at the university make me feel at home.	50	***62***
10.	I feel that I fit in well at the university.	68	***82***

Test of Young Adult Social Skills Knowledge. Table 13 shows that Chase's score on the Test of Young Adult Social Skills Knowledge (TYASSK) increased modestly from 16 out of a possible 30 at pre-test to 17 out of 30 at post-test. This change indicated a slight gain in his social skills knowledge following the intervention. Although this single-point improvement suggested some assimilation of the curricular content, the relatively small magnitude of change likely reflected both the abbreviated and modified nature of the treatment phase implemented in this study. Furthermore, it was noted that several of Chase's incorrect responses at pre- and post-test pertained to topics (e.g., relationship management, direct bully confrontation) that were not covered in depth during the modified intervention due to resource and time constraints.

Table 13. Chase Pre-Post Test of Young Adult Social Skills Knowledge

Participant	Pre	Post
- Chase	16	17

Notes: *Many incorrect responses correlated with sections of the curriculum we did not exclusively cover (i.e., relationships, direct bully confrontation). This requires further exploration in future research, especially since the treatment phase was substantially shorter and heavily modified compared to the recommendations.*

Social Validity. Chase's ratings on the CIRP-M indicated very high social validity for the program (See Table 14). Across ten items, his average endorsement was 89%, with many statements,

such as "The facilitators were supportive" and "I would recommend this program to other students," receiving perfect or near-perfect scores (99–100%). He also rated the program as fair (87%), enjoyable (93%), and useful (92%), and agreed that it addressed issues important to him (100%) and benefited his university life (91%). The lowest endorsements were for statements concerning concrete gains, such as "The program helped me improve my social skills" (74%) and "I feel more confident in social situations now" (75%). This aligned with his comment that he had already received prior social-skills training and therefore saw the greatest utility for peers who lacked that background. Overall, these data suggested that Chase found the program acceptable, relevant, and beneficial, particularly as a model for others, even if his perceived personal gains in skill acquisition were somewhat tempered by his existing skill level.

Table 14. Chase Pre-Post (Children's Intervention Rating Profile–Modified)

	Statement	Score
1.	The program was fair.	87
2.	I enjoyed participating in this program.	93
3.	The facilitators were supportive.	99
4.	The strategies taught were useful to me.	92
5.	The program helped me improve my social skills.	74
6.	I feel more confident in social situations now.	75
7.	The program addressed issues important to me.	100
8.	I would recommend this program to other students.	100
9.	Participating in the program benefited my university life.	91
10.	The program was a good way to address my social challenges.	82

Comments: *"I feel this program could help others a lot more than me. I had previously received help with developing my social skills, but I understand not everyone has had that opportunity. This program could be extremely beneficial to others."*

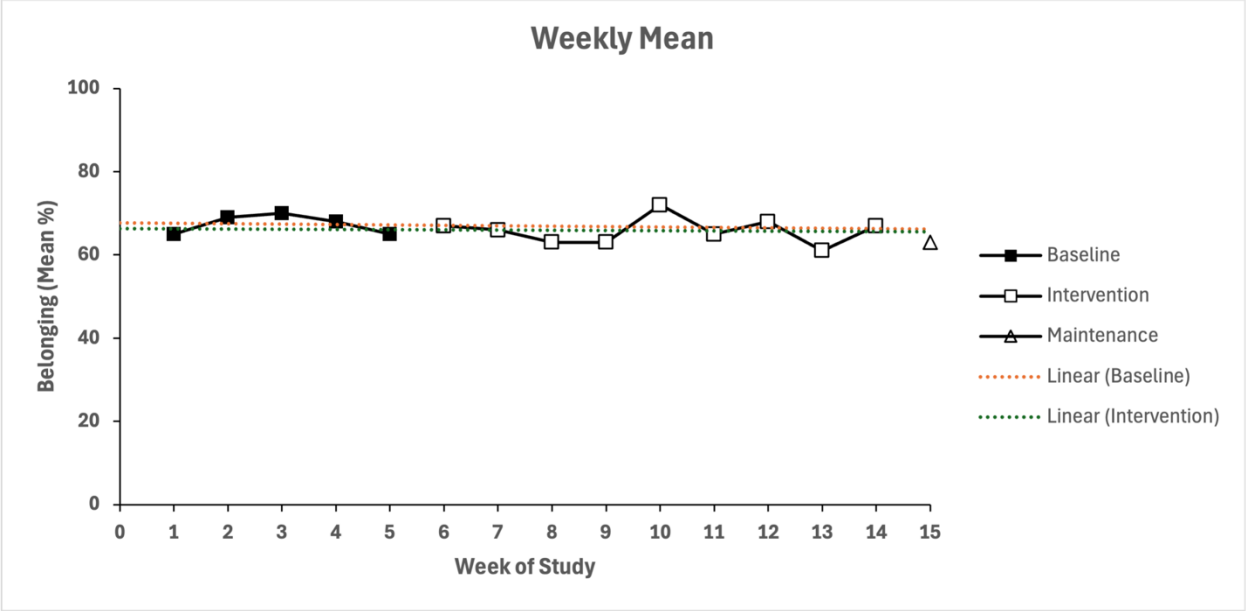
Alyx

Belonging

Preliminary Visual Analysis. Visual inspection of Alyx's weekly mean belonging scores (See Figure 7), applying the six Ledford et al. (2017) criteria, revealed a response to supports that was characterized by considerable overlap between phases. During the initial five-week baseline, Alyx's scores remained relatively constant, exhibiting minimal variability around a slight peak in

Week 3 before a modest downturn by Week 5. The introduction of supports in Week 6 did not yield an immediate positive change; instead, mean belonging showed an initial stasis followed by a slight dip in Weeks 8 and 9, suggesting a delayed therapeutic response. A notable upward spike occurred in Week 10, after which scores fluctuated around the intervention's mean for the remainder of the study. Upon withdrawal of supports, the single data point recorded for the maintenance phase was positioned just below the baseline envelope, making definitive interpretation of a sustained effect challenging due to insufficient data. In summary, while the distinct peak in Week 10 offered strong visual evidence of an intervention-related effect, the substantial data point overlap observed from Weeks 6 to 9 and again from Weeks 11 to 13 indicated that Alyx's overall shift in mean belonging was more incremental and less sharply defined than a pronounced phase-to-phase level change.

Figure 7. Alyx's Belonging (Weekly Mean)

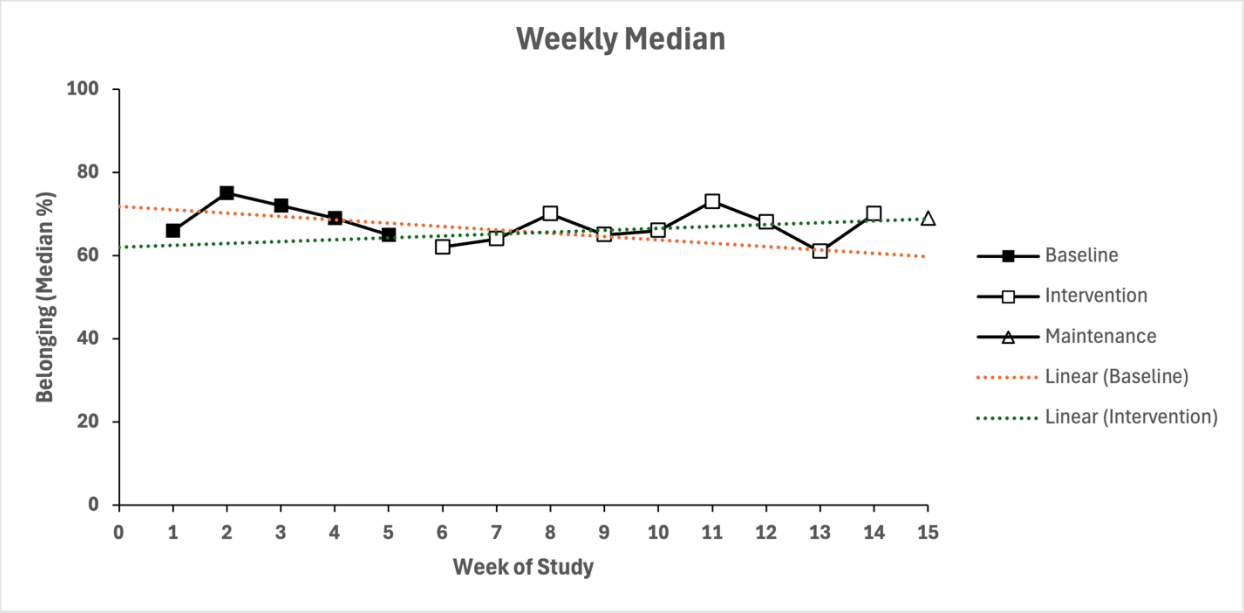


Alyx's weekly median belonging scores (See Figure 8), evaluated against the six visual-analysis criteria (Ledford et al., 2017), painted a picture of even more gradual change and extensive inter-phase overlap when compared to the mean-based analysis. The baseline phase for median scores displayed moderate variability, with a slight downward trajectory noted after Week 2. When

supports were initiated in Week 6, median belonging first declined that week, then rose by Week 7, but without an immediate shift in level above the baseline range; these early intervention medians were situated at or just below the lower boundary of the baseline envelope. In Week 8, the median climbed briefly before returning to baseline levels. A more sustained upward movement did not materialize until Week 11, when the median neared, but did not surpass, the highest baseline value.

Critically, across the nine intervention data points, eight were contained within the baseline envelope, demonstrating pronounced overlap in level despite an emerging upward trend late in this phase. Variability during the intervention remained moderate as medians oscillated. Following the discontinuation of supports in Week 14, the solitary maintenance median once more fell within the baseline range, precluding firm conclusions about a maintenance plateau due to the single data point. To summarize, although a gradual therapeutic trend culminated in Week 11, suggesting a delayed intervention effect, the pronounced overlap throughout much of the intervention phase (Weeks 6 to 9 and 10 to 13) signified that Alyx's shift in median belonging lacked the immediacy and distinctness of a classic phase-differentiated level change. The mean-based analysis, by contrast, had detected an earlier shift at Week 10; this suggested the mean may be more reactive to singular significant data points, whereas the median more robustly reflected Alyx's central tendency.

Figure 8. Alyx's Belonging (Weekly Median)



When plotted across all daily prompts (n = 151; See Appendix B: Figure 5), Alyx's raw belonging scores exhibit much greater moment-to-moment variability than the weekly summaries, yet they nonetheless trace the same phase-linked patterns of change. During the baseline phase, daily scores swung widely, with no clear sustained trend. Instead, values regularly oscillated around the mean, producing substantial within-phase variability and masking any incipient level or slope shift prior to the supports. As soon as intervention began, Alyx's reported scores continued to fluctuate around the mean, with noticeable spikes at certain intervals. However, these high scores were punctuated throughout the intervention phase by intermittent dips well below the mean, highlighting fluctuations that average out in the aggregated weekly plots. While the weekly mean captured the phase change the moment a cluster of high-score days inflated the phase average, and the weekly median provided a more outlier-resistant confirmation after multiple high days accumulated, the daily series shows the sudden leaps and temporary reversals that belie an immediacy of effect. When supports were withdrawn, there was an immediate, sharp downtown before rebounding. In sum, although the daily plot reveals erratic "spikes" and "troughs" that are smoothed by weekly

aggregation, the same phase-to-phase increases, reversals, and replications appear: baseline instability gives way to a cluster of higher scores under intervention, and then to a steep decline once supports end, confirming that weekly means or medians faithfully capture the intervention's overall effect while filtering out the significant day-to-day variability.

Structured Visual Analysis. The CDC analysis applied to Alyx's weekly belonging scores (See Table 15) failed to establish a statistically reliable intervention effect, irrespective of whether mean or median summaries were examined. Specifically, the assessment of weekly mean data showed that only a single treatment-phase week (one of nine) exceeded both therapeutically adjusted criterion lines. The weekly median data presented an even more conservative picture, with no treatment-phase medians (zero of nine) meeting these dual criteria. As the binomial test ($p < .05$) stipulated a minimum of eight such "successes" for statistical significance, the CDC outcome for Alyx's belonging remained non-significant for both aggregation methods. This difference between one success (mean) and zero successes (median) was minimal and reinforced the interpretation that occasional peaks in Alyx's belonging ratings, like the one observed in Week 10, did not occur with sufficient consistency to satisfy the CDC's rigorous binomial requirement. An attempt to apply the CDC methodology to Alyx's entire daily belonging dataset was not undertaken due to the substantial re-engineering required for the existing spreadsheet macros to handle the increased data load. For subsequent research, the plan involves developing enhanced computational tools for this purpose, potentially by overhauling the current macros or by employing more powerful statistical software (e.g., *R*), which will permit a comprehensive CDC evaluation of the full daily data series.

Table 15. Alyx's Conservative Dual-Criterion (Belonging)

Weekly	# of Treatment Points	# Above Both Lines	# Needed Above Both Lines
- Mean	9	1	8
- Median	9	0	8

Non-Overlap of All Pairs and Within-Case Quantitative Index. For Alyx, the transition from baseline to intervention (See Table 16) resulted in a NAP of .40 (95% CI [.36, .55]), a value signifying that fewer than half of their intervention-phase belonging scores exceeded those recorded during baseline. The corresponding LRR was -0.03 (95% CI [-0.10, 0.04]), which reflected an approximate 3% proportional decrease (95% CI [-10%, 4%]) in mean belonging after supports were introduced. When supports were withdrawn and the maintenance phase commenced, the NAP value comparing intervention to maintenance showed only a marginal increase to .50 (95% CI [.32, .67]), meaning just half the maintenance scores surpassed intervention-phase values; the LRR for this transition was -0.04 (95% CI [-0.26, 0.19]), equivalent to a 3% mean decline (95% CI [-23%, 20%]). Considering that both LRR confidence intervals encompassed zero and neither this phase shift nor the baseline-to-intervention transition had yielded the eight "successes" stipulated by the CDC method for statistical significance, these NAP and LRR results collectively indicated that Alyx's sense of belonging neither improved substantially during the intervention nor showed a marked rebound in the maintenance phase. In summary, the supports did not appear to produce a reliable or meaningful gain in Alyx's belonging, with maintenance-phase levels remaining essentially comparable to, or slightly below, those observed during the intervention.

Table 16. Alyx's Belonging (NAP; LRR)

Domain	Phase Shift	NAP [95% CI]	LRR [95% CI]	% Change [95% CI]
- Belonging	BL>>>INT	0.45 [0.36, 0.55]	-0.03 [-0.10, 0.04]	-3 [-10, 4]
- Belonging	INT>>>MN	0.50 [0.32, 0.67]	-0.04 [-0.26, 0.19]	-3 [-23, 20]

Confounding Variable: Social Partner. Table 17 presents Alyx's self-reported belonging scores, analyzed according to whether a social partner was present during each data prompt. A consistent and notable pattern emerged: across all three study phases, the presence of a social partner was linked with substantially higher belonging scores. For instance, during the baseline period, Alyx's belonging scores averaged 75.2 when with a social partner, compared to 56.6 when

alone, a substantial difference of nearly 19 points. This trend persisted into the intervention phase, where ratings remained higher when accompanied by a partner (M = 73.0) versus when reported alone (M = 58.8), still reflecting a considerable gap of over 14 points. The disparity became even more marked in the maintenance phase, with mean belonging reaching 74.0 when a social partner was present, versus only 52.0 when Alyx was not accompanied. This overarching pattern was mirrored in the median scores across all phases. Variability in scores was generally greater during instances when a social partner was present, likely attributable to a wider array of contextual social interactions experienced. These persistent and sizable differences suggested that the presence of a social partner potentially acted as a significant influential factor, and a potential confound, on Alyx's self-perceived belonging throughout all conditions of this study.

Table 17. Alyx's Social Partner Report (Belonging)

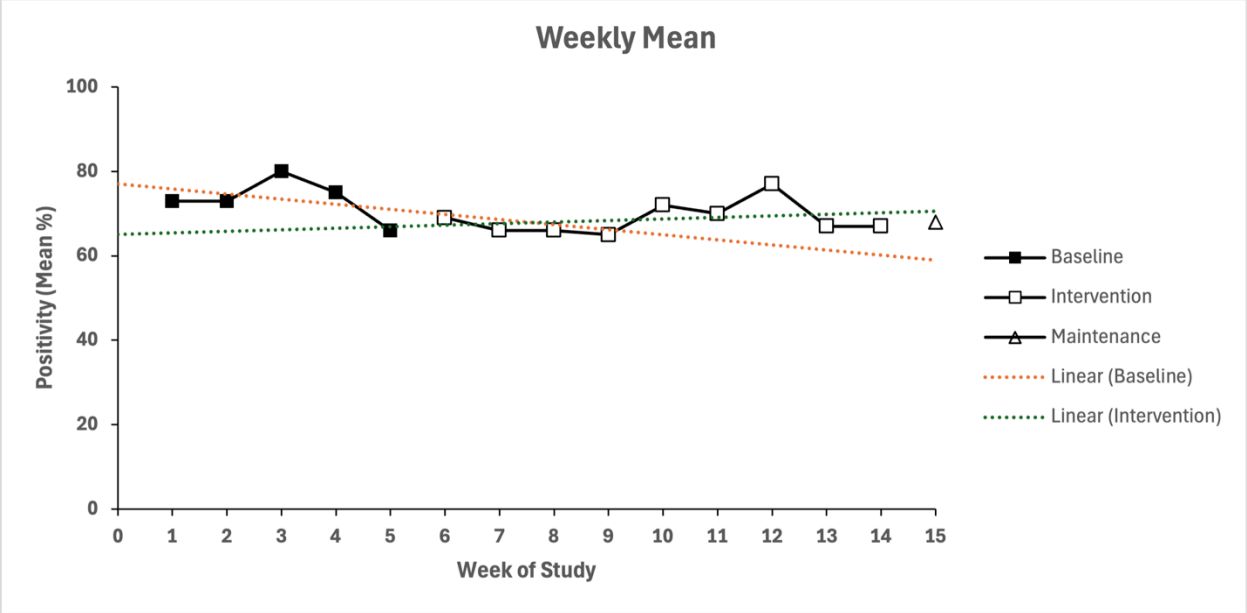
Condition	Social Partner	N	Mean	Median	SD	Min	Max
- Baseline	Y	32	75.2	76.0	12.03	45	100
- Baseline	N	22	56.6	55.0	10.29	35	80
- Intervention	Y	42	73.0	74.0	14.92	30	100
- Intervention	N	45	58.8	60.0	9.06	39	79
- Maintenance	Y	5	74.0	82.0	19.51	41	88
- Maintenance	N	5	52.0	56.0	20.06	23	74

Positivity

Preliminary Visual Analysis. An assessment of Alyx's weekly mean positivity scores (See Figure 9) using the six visual-analysis criteria (Ledford et al., 2017) indicated a more nuanced intervention effect compared to that observed for their belonging, with substantial overlap evident across phases. The initial baseline period was marked by moderate variability and a slight downward trajectory in mean positivity, declining from a peak in Week 3 to a low point by Week 5. Following the introduction of supports in Week 6, mean positivity initially decreased modestly, remaining within the baseline envelope, and subsequently declined further by Week 9; this pattern suggested a minimal immediacy of effect. Commencing in Week 10, however, the intervention appeared to

stimulate a gradual upward trend in scores, which persisted until peaking in Week 12. Despite this mid-intervention upward shift, a high degree of overlap was present, as eight of the eight intervention data points fell within the baseline range. Variability during the intervention phase remained moderate, with values showing narrow oscillations around the mean. With the withdrawal of support in Week 14, the single maintenance score continued to reside within the original baseline envelope, precluding definitive conclusions about sustained effects from this solitary data point. In summary, although a gradual therapeutic trend from Week 10 onwards provided some visual support for an intervention effect, fulfilling criteria for trend and level change, the pronounced overlap during Weeks 6 to 9 and Weeks 11 to 14, alongside moderate variability, suggested that Alyx's improvement in mean positivity was incremental and less distinct than a clear, phase-differentiated change.

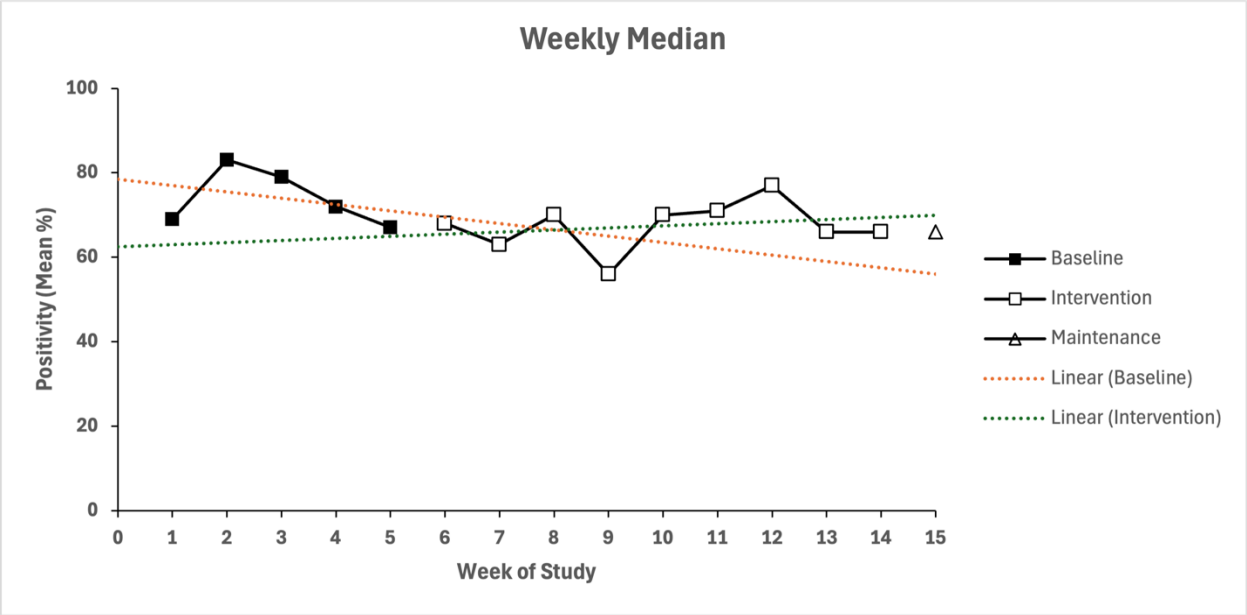
Figure 9. Alyx's Positivity (Weekly Mean)



Alyx's weekly median positivity scores (See Figure 10), when examined through the lens of Ledford et al.'s (2017) six visual-analysis criteria, presented a pattern subtly distinct from the mean-based analysis, primarily characterized by a more delayed response and greater sensitivity to lower

scores. During baseline, median positivity, while relatively high and peaking in Week 2, exhibited a slight downward slope by Week 5 amidst modest variability. The introduction of supports in Week 6 saw median positivity dip initially, remaining within the baseline envelope, and then fall further in Week 7, indicative of a delayed and minimal change in level. A brief rebound occurred in Week 8, only for the median to drop again in Week 9, reaching its lowest intervention value. Thereafter, a climb in medians was observed by Week 13. Intervention phase variability was notably greater than that in baseline, and unlike the mean series, no single week displayed dramatic upward surges. Upon withdrawal of support in Week 14, the maintenance median once again overlapped with the baseline envelope, rendering it challenging to infer maintenance of gains from this single observation. The mean-based analysis had registered a gradual upward trend beginning somewhat earlier (Week 10) and had suggested a clearer immediacy of effect upon that trend's emergence. The median scores, in contrast, lagged in indicating positive change and appeared more susceptible to the influence of intermittent low-scoring weeks (e.g., Week 9). This comparison indicated that for Alyx's positivity, the mean tended to smooth data into an incremental increase, whereas the median captured a more punctuated pattern of dips and rebounds, being more reflective of the data's central tendency when influenced by these lower outlying weeks and thus showing a more delayed overall upward shift.

Figure 10. Alyx's Positivity (Weekly Median)



Examination of Alyx's complete set of daily positivity ratings (n = 151; See Appendix B: Figure 6) revealed that scores fluctuated widely yet generally followed the core phase-linked patterns observed in the weekly summaries. During the baseline phase, the daily values exhibited no discernible, sustained trend; instead, they demonstrated substantial within-phase variability, often oscillating around the mean, which effectively masked any subtle, gradual shifts in level or slope before supports were introduced. Once the intervention began, several of Alyx's daily scores showed abrupt leaps into the upper range, mirroring the Week 10 surge noted in the weekly mean analysis. However, these peaks were interspersed with intermittent drops into the lower range on other days. While the weekly mean analysis had captured the intervention effect once a cluster of high-scoring days influenced the average, and the weekly median had subsequently confirmed that effect with its outlier resistance, the daily series uniquely highlighted both pronounced spikes (e.g., several "100%" days) and significant troughs (e.g., around 50%) that were averaged out in the weekly aggregations. Following the withdrawal of supports, the initial maintenance reading showed a sharp decline, which was then followed by a rapid rebound later in the maintenance period; this daily pattern precisely

mirrored the steep decline and partial recovery also registered in both weekly median and mean summaries. In essence, although Alyx's daily plot displayed considerable moment-to-moment volatility with erratic "spikes" and "troughs" not apparent in the weekly charts, the fundamental phase-to-phase increases, reversals, and replications still emerged. This correspondence confirmed that aggregating to weekly means or medians appropriately captured the intervention's overall directional influence for Alyx, while effectively filtering out their substantial day-to-day response variability.

Structured Visual Analysis. Application of the CDC method to Alyx's weekly positivity scores, for both mean and median summaries, did not indicate a statistically reliable intervention effect (See Table 18). Specifically, neither the analysis of weekly means nor weekly medians yielded any treatment-phase data points (zero out of nine for both) that surpassed the two stringently defined criterion lines. Given that eight such "successes" were necessary to meet the binomial threshold for statistical significance ($p < .05$), the CDC criteria were not met for Alyx's positivity data using either aggregation method. This consistent null finding across both mean and median calculations underscored that, despite occasional high positivity values observed during the intervention phase (e.g., a peak in Week 12), these instances were not frequent enough to satisfy the strict requirements of the binomial test. Extending the CDC procedure to Alyx's complete daily positivity dataset was deemed unfeasible with the existing spreadsheet macros, as this would have necessitated a significant recalibration of the embedded formulas to manage the larger data volume. Future work will involve adapting these macros or implementing the CDC computation within a statistical package (e.g., *R*), to make a full-data assessment possible.

Table 18. Alyx's Conservative Dual-Criterion (Positivity)

Weekly	# of Treatment Points	# Above Both Lines	# Needed Above Both Lines
- Mean	9	0	8
- Median	9	0	8

Non-Overlap of All Pairs and Within-Case Quantitative Index. For Alyx, the transition from baseline to intervention (See Table 19) yielded a NAP of .40 (95% CI [.31, .50]), a value indicating that only 40% of their intervention-phase positivity scores exceeded those recorded during baseline. The corresponding LRR was -0.07 (95% CI [-0.14, 0.00]), which reflected an approximate 7% proportional decrease (95% CI [-13%, 0%]) in mean positivity after supports were introduced. When supports were withdrawn and the maintenance phase commenced, the NAP value comparing intervention to maintenance rose only slightly to .47 (95% CI [.30, .65]), signifying that fewer than half of the maintenance-phase scores surpassed intervention-phase values. The LRR for this transition was -0.01 (95% CI [-0.19, 0.17]), equivalent to a negligible 1% mean decline (95% CI [-18%, 18%]). Given that both LRR confidence intervals included zero and the NAP values fell substantially below conventional thresholds for a reliable effect, these findings collectively suggested that Alyx's positivity neither increased meaningfully during the intervention nor rebounded significantly in the maintenance phase. In essence, the supports did not appear to generate a statistically or practically significant improvement in their positivity; maintenance-phase levels consequently remained largely comparable to, or slightly below, those observed during the intervention phase.

Table 19. Alyx's Positivity (NAP; LRR)

Domain	Phase Shift	NAP [95% CI]	LRR [95% CI]	% Change [95% CI]
- Positivity	BL>>>INT	0.40 [0.31, 0.50]	-0.07 [-0.14, 0.00]	-7 [-13, 0]
- Positivity	INT>>>MN	0.47 [0.30, 0.65]	-0.01 [-0.19, 0.17]	-1 [-18,18]

Confounding Variable: Social Partner. An analysis of Alyx's positivity scores, conditional upon social partner presence (See Table 20), revealed a consistent association between being with others and higher reported positivity across all study phases. During the baseline period, a notable disparity was evident: mean positivity when a social partner was present (M = 79.1) contrasted sharply with scores when Alyx was alone (M = 65.5), a difference of nearly 14 points. This tendency

for elevated positivity in social contexts persisted through the intervention phase, where mean scores with a partner (M = 72.5) remained higher than those recorded when alone (M = 65.2). The influence of social context appeared even more pronounced during the maintenance phase, as positivity ratings when Alyx was with a partner (M = 76.6) were markedly greater than when they were not accompanied (M = 58.6), an 18-point difference. Median scores corroborated this overarching trend across all phases, and score variability tended to be greater when a social partner was present, likely reflecting a broader spectrum of emotional experiences within social interactions. The persistence and magnitude of these differences strongly suggested that social partner presence was a key determinant of Alyx's reported positivity levels. Consequently, this factor potentially functioned as a significant confounding variable when attempting to isolate the specific effects of the intervention on their positivity. Future research should consider methods for controlling social context or gathering additional qualitative data to further elucidate how moment-to-moment peer interactions influence emotional self-reporting among neurodiverse students.

Table 20. Alyx's Social Partner Report (Positivity)

Condition	Social Partner	N	Mean	Median	SD	Min	Max
- Baseline	Y	32	79.1	79.0	12.3	48	100
- Baseline	N	22	65.5	65.5	14.4	48	100
- Intervention	Y	42	72.5	71.0	14.8	45	100
- Intervention	N	45	65.2	65.0	10.8	45	85
- Maintenance	Y	5	76.6	79.0	18.2	49	100
- Maintenance	N	5	58.6	54.0	17.0	41	87

Measurement

Social Connectedness Scale—Revised (Modified). As Table 21 illustrates, Alyx's pre- and post-intervention scores on the SCSR-M demonstrated substantial gains in their overall sense of campus connectedness. Their average score across the ten items shifted from 43.9 at baseline to 67.7 at post-test, an overall mean increase of 23.8 points. Meaningful improvement was evident on nearly all individual items. The most considerable advancement occurred for the item "I feel comfortable

interacting with other students and professors I don't know well," which surged from a score of 0 to 49 (+49 points). Other significant increases were noted in perceived approachability and friendliness of others ("I see other students and professors as approachable and friendly," from 44 to 75, +31 points), ease of connecting ("I find it easy to connect with other students and professors," from 15 to 44, +29 points), and general academic/social belonging ("I feel like I belong in social and academic settings at the university," from 38 to 66, +28 points). Further positive changes were registered for feelings of connection to the broader university community (+27 points), a sense of fitting in at the university (+21 points), and feeling at home with university friends (+21 points).

In contrast to these widespread gains, Alyx reported a slight decrease on one item: "I feel understood and accepted by my peers and professors" (from 85 to 77, -8 points). This particular finding suggested a nuanced outcome where, although their overall connectivity and comfort in initiating interactions improved, their perception of being understood and accepted by peers and faculty did not follow the same positive trajectory. Collectively, these results indicated that the intervention markedly enhanced Alyx's social connectedness, particularly in areas related to comfort with new interactions, their view of others' approachability, and a broader sense of community belonging. Simultaneously, the data highlighted that feelings concerning peer and faculty acceptance might represent a distinct area warranting additional support or future research exploration.

Table 21. Alyx Pre-Post (Revised Social Connectedness Scale–Modified)

	Question	Pre	Post
1.	I feel connected to other students on campus.	75	***78***
2.	I actively participate in campus groups or activities.	70	***81***
3.	I feel like I belong in social and academic settings at the university.	38	***66***
4.	I find it easy to connect with other students and professors.	15	***44***
5.	I feel understood and accepted by my peers and professors.	***85***	77
6.	I see other students and professors as approachable and friendly.	44	***75***
7.	I feel comfortable interacting with other students and professors I don't know well.	0	***49***
8.	I feel connected to the broader university community.	30	***57***
9.	The friends I have made at the university make me feel at home.	69	***90***
10.	I feel that I fit in well at the university.	39	***60***

Test of Young Adult Social Skills Knowledge. Table 22 shows that Alyx's performance on the TYASSK improved notably from a pre-test score of 16 (out of a possible 30) to 23 out of 30 at post-test. This seven-point increase reflected a substantial gain in their social skills knowledge following the intervention and suggested a meaningful assimilation of the material covered. This improvement was particularly noteworthy considering that several sections of the TYASSK, specifically those pertaining to relationship management and direct bullying response strategies, were not part of the modified curriculum used in this study. Consequently, it is likely that incorrect responses on these omitted sections constrained Alyx's potential maximum score, implying that their knowledge acquisition in the domains actually taught may have been even greater. The abbreviated and modified nature of the treatment phase in this dissertation study further underscored the importance for subsequent research to explore how variations in curriculum fidelity and intervention length impact the extent of social skills knowledge acquisition.

Table 22. Alyx Pre-Post Test of Young Adult Social Skills Knowledge

Participant	Pre	Post
- Alyx	16	23

Notes: *Many incorrect responses correlated with sections of the curriculum we did not exclusively cover (i.e., relationships, direct bully confrontation). This requires further exploration in future research, especially since the treatment phase was substantially shorter and heavily modified compared to the recommendations.*

Social Validity. Table 23 presents Alyx's social validity ratings on the CIRP-M, which revealed generally positive perceptions of the program alongside some areas of tempered enthusiasm. Alyx assigned perfect scores (100%) to statements concerning program fairness, facilitator support, and their willingness to recommend the program, indicating a strong endorsement of these fundamental aspects. High ratings were also given for perceived benefits to university life (91%) and the belief that the program constituted a good approach to addressing social challenges (85%). In contrast, Alyx's rating for enjoyment of participation was 88% and for

the usefulness of strategies was 84%, suggesting solid, albeit not uniformly high, engagement with the content.

Lower endorsements were evident on items assessing concrete personal outcomes: "The program helped me improve my social skills" (72%), "I feel more confident in social situations now" (67%), and "The program addressed issues important to me" (68%). These more moderate ratings may have reflected Alyx's existing strengths or perhaps specific expectations regarding curriculum components that were not fully met. Alyx also provided a non-verbal comment in the form of a smiley face emoji (😊), which conveyed an overall positive sentiment, even though endorsements for some efficacy-related items were not at the maximum level. Collectively, these results suggested that while Alyx found the program to be fair, supportive, and recommendable, their self-reported gains in social skill development and confidence were modest. This outcome highlighted potential opportunities to further tailor program content or delivery to individual participant needs and expectations in future iterations.

Table 23. Alyx Pre-Post (Children's Intervention Rating Profile – Modified)

Statement	Score
1. The program was fair.	100
2. I enjoyed participating in this program.	88
3. The facilitators were supportive.	100
4. The strategies taught were useful to me.	84
5. The program helped me improve my social skills.	72
6. I feel more confident in social situations now.	67
7. The program addressed issues important to me.	68
8. I would recommend this program to other students.	100
9. Participating in the program benefited my university life.	91
10. The program was a good way to address my social challenges.	85

Comments: "😊"

Discussion

This SCED study aimed to evaluate the effects of a heavily modified PEERS®-YA curriculum on social engagement, specifically how it affected the sense of belonging and overall

feelings of positivity for autistic undergraduates at a midwestern university. The investigation additionally sought to assess changes in participants' social skills knowledge and gather their perceptions of the intervention's social validity. In this chapter, I will synthesize the findings, interpret their significance in the context of extant literature, discuss methodological considerations, including the inconclusiveness of overall effects, compare participant outcomes, address potential confounding variables, and propose directions for practice and future research.

Summary and Interpretation of Major Findings: An Inconclusive Picture

My primary conclusion drawn from the comprehensive triangulation of data, encompassing preliminary visual analysis of time-series data, statistical corroboration via the CDC method, non-overlap and effect size calculations (i.e., NAP, LRR), and pre-post measures, is that the effects of the modified PEERS®-YA intervention implemented in this study were largely inconclusive regarding its consistent efficacy across participants. While some positive changes in belonging and positivity were visually apparent for one participant (Chase), these were not consistently replicated with the same clarity or magnitude for the other participant (Alyx). Furthermore, statistical authentication using the stringent CDC method was not achieved for either participant on their weekly mean and median reported outcomes of belonging or positivity, with unknown results when assessing the full daily reported data set (n = 178 for Chase; n = 151 for Alyx).

Conservative Dual-Criterion. The consistent lack of statistically significant findings using the CDC method in this study, despite some visually apparent positive trends or level changes, warrants careful consideration. Adopted by Fisher et al. (2003), the CDC method is well-regarded for its formulaic approach to visual analysis in SCEDs and is noted for strengths in controlling Type I errors (i.e., false positives). This is achieved by requiring data in the intervention phase to surpass two criteria derived from baseline data, a mean line and a trend line adjusted by 0.25 standard deviations in the predicted direction of behavior change, for a prespecified number of observations.

However, this inherent stringency, while beneficial for reducing the likelihood of incorrectly claiming an effect, can simultaneously increase the risk of Type II errors (i.e., false negatives). Consequently, a true intervention effect may be overlooked, a concern particularly pertinent in applied research characterized by small sample sizes, shorter data series, or inherent data variability (Fisher et al., 2003).

In the current study, I applied the CDC method to weekly aggregated data for both participants, which required a minimum of eight data points in the intervention phase to fall beyond both the adjusted mean and trend lines to be deemed statistically significant. For Chase, visual inspection of weekly mean and median belonging data suggested a positive intervention effect, with seven out of nine intervention data points meeting these demanding criteria. Despite this strong visual trend, the predetermined threshold for CDC significance was narrowly missed. This outcome highlights a key limitation of the CDC method, being its binary nature to label results as "significant" or "not significant," which may not fully capture the nuances of behavioral change, particularly when effects are moderate, or data are bordering the decision threshold. Fisher et al. (2003) noted that, while the CDC method maintains reasonable statistical power compared to some statistical alternatives in shorter data series (e.g., dual criteria method), the primary benefit lies in conservative control of Type I errors (i.e., false positives).

For Alyx, whose data for both belonging and positivity exhibited greater variability and fewer distinct changes between phases, the CDC criteria were consistently unmet. This underscores another aspect of the CDC method, being the performance may be influenced by the characteristics of the data. While the CDC method offers a simpler way to protect against the biasing effects of serial dependence compared to more complex statistical models, high variability or less pronounced effects can potentially limit the ability to meet the rigid criteria.

Regardless, the utility of the CDC method as an objective benchmark is a significant benefit since it aims to improve the accuracy, objectivity, and interrater reliability of visual inspection (Fisher et al., 2003; Swoboda et al., 2010) and has shown to foster strong agreement with expert visual analysts when experts concur (Wolfe et al., 2018). Furthermore, it has proven effective as a training tool for teaching visual inspection skills (Fisher et al., 2003; Stewart et al., 2007).

However, several limitations beyond the CDC method's conservative nature are relevant. In the original formulation, the CDC method primarily focuses on identifying systematic change rather than the magnitude or clinical significance of an effect (Fisher et al., 2003; Swoboda et al., 2010). An experienced visual inspector might discern a clinically meaningful change that the CDC method, due to numerical criteria, does not flag as statistically significant. Additionally, the reliance on visual aids (i.e., plotted criterion lines, binomial tables) for accurate implementation may be a drawback, specifically when analysts must apply these skills to published literature when visual aids are typically absent from graphs (Stewart et al., 2007). The original statistical underpinnings, specifically the binomial test employed, is noted for not fully accounting for the autocorrelated nature of much single-case data, although Fisher et al. (2003) demonstrated the robustness to certain levels of autocorrelation. Extensions of the CDC method to more complex designs (e.g., reversal, multiple baseline) have been proposed but might be considered informal, requiring further empirical validation (Swoboda et al., 2010).

In the context of this study, the CDC method provided a standardized and objective framework for data analysis. However, its conservative nature, coupled with the specific data patterns observed, particularly the near-miss for Chase and the higher variability for Alyx, likely contributed to the overall inconclusive statistical findings regarding intervention efficacy on the primary time-series measures. While valuable for its rigor and control of false positives, the CDC method's limitations, especially its potential for Type II errors and its focus on statistical rather than

clinical significance, must be acknowledged when interpreting outcomes in applied settings with nuanced data. Future research might explore complementary analyses that can capture effect magnitude and social validity alongside the systematic change detection offered by the CDC method.

Comparison of Participant Outcomes

There was significant heterogeneity in the response to intervention for my study. While Chase typically demonstrated more consistent and visually apparent positive responses on the daily ESM measures of belonging and positivity, the clear phase changes present in weekly data did not meet CDC criteria, although NAP and LRR values indicated improvements during the intervention phase, albeit modest to moderate shifts. Furthermore, his maintenance data indicated potential retention of these gains above baseline levels; however, Chase's TYASSK score showed a marginal 1-point gain, with several of his incorrect responses pertaining to topics (e.g., relationship management, dating skills, direct bully confrontation) not explicitly covered in the heavy modified curriculum. Despite the limited gain in assessed knowledge, his social validity ratings were quite high, suggesting he found the program fair, supportive, and recommendable, potentially showing value in the structured interactions and supportive environment more than direct skill acquisition given his prior experience with developing some of these skills.

Alyx's response profile was markedly different, with their weekly belonging and positivity scores characterized by considerable overlap between phases, more gradual or delayed changes, and greater moment-to-moment variability in the ESM Data. NAP and LRR values for both belonging and positivity during the baseline-to-intervention transition did not indicate improvement, although the visual analysis of their mean positivity did reveal a delayed therapeutic trend starting mid-intervention; however, the overall interpretation of results was less clear and more variable than Chase's. In contrast to their daily ESM data, and to Chase's results, Alyx demonstrated a substantial 7-point gain on the TYASSK, suggesting a meaningful assimilation of the taught material, even with

the significant modifications. Furthermore, their social validity ratings were generally positive, especially regarding program fairness and facilitator support.

These contrasting trajectories underscore the well-documented heterogeneity among autistic individuals. This suggests that responses to structured interventions like PEERS®-YA can vary significantly based on individual characteristics (e.g., baseline levels, pre-existing skills, co-occurring conditions), how they engage with specific intervention components, and potentially factors that went unmeasured in this study.

Extant Literature Connection

Several findings warrant further discussion in relation to existing literature.

Social Partner Presence

While not a primary variable, a particularly noteworthy finding included the potential for social partners to influence participant outcomes, especially for Alyx, who reported a strong, consistent association between the presence of a social partner and higher self-reported belonging and positivity across all study phases. This indicates that, for Alyx, the immediate social context is potentially a powerful moderator of their reported internalized states. While this observation is valuable, highlighting the importance of social engagement for overall wellbeing, it simultaneously introduced a significant potential confound when interpreting the intervention's direct effects on their outcomes.

While the intervention aimed to build skills that could be applied to initiate or improve social interactions, Alyx's higher scored predominantly occurred when they reported engaging with a social partner, which complicates attributing these higher scores solely to the intervention's skill-building components. These results align with literature indicating that the availability and quality of peer support are crucial for the well-being and adjustment of autistic university students (Anderson et al., 2017; Schembri-Mutch et al., 2023). Additionally, it emphasized the methodological challenge of

conducting SCED research in naturalistic settings and extracting intervention effects from naturally occurring variations in social opportunities. Furthermore, Alyx's self-reported data when including a social partner did include some significant dips, suggesting that the specific social partner (e.g., peer, professor, sibling, parent) may have a negative impact on their overall sense of belonging and feelings of positivity, necessitating further inquiry into the specific impact by reported individual rather than simply being socially engaged.

Pre-Post Measure Changes

Both participants reported overall gains on the SCSR-M at the culmination of the study, with Chase moving from a mean of 60.0 to 77.1, and Alyx from 43.9 to 67.7. Given the significant variability and less predictable nature of Alyx's ESM data, their overall sense of connectedness and belonging on campus saw significant improvement. This discrepancy may indicate that, while momentary feelings of belonging were variable and not consistently improved by the intervention for Alyx, their global perception of connectedness to the campus environment was enhanced throughout the semester.

The TYASSK results presented an inverse pattern, with Alyx showing a larger gain in social skills knowledge, despite the higher variability in ESM changes, than Chase, who had more visually positive ESM trends but a smaller TYASSK gain. Chase's comments of prior training and the modified curriculum's omissions may account for his modest TYASSK change; whereas Alyx's stronger TYASSK gain, despite the ESM outcomes, could imply a genuine acquisition of developing knowledge that had not consistently translated into changed daily experiences of self-perceptions or ESM did not fully capture the domains of increased competencies.

Social Validity

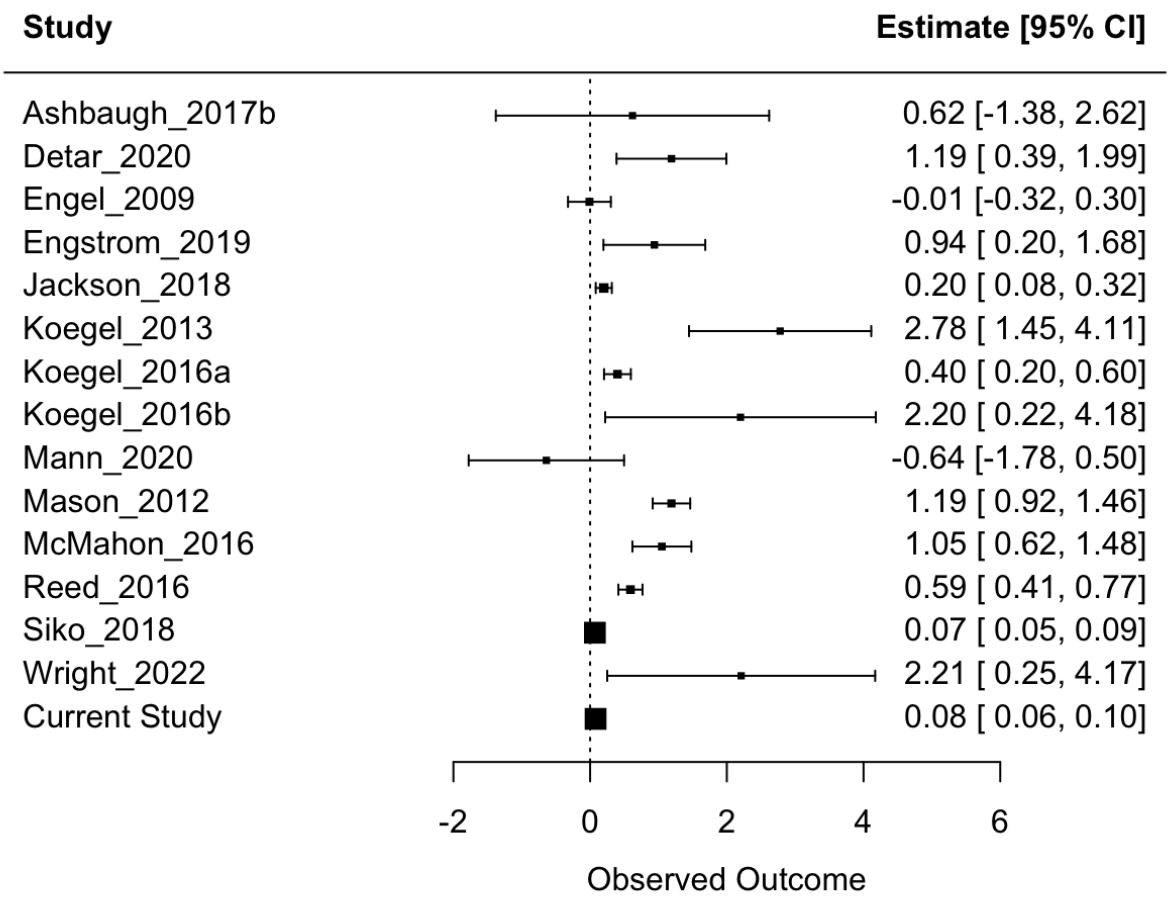
Both participants reported high social validity ratings, especially regarding the program fairness, facilitator support, and suitability, with both indicating they would recommend the

intervention to their peers. Chase, despite minimal TYASSK gains, valued the program highly. Alyx, while reporting more modest personal gains in confidence and skills on the CIRP-M, still provided positive feedback, including a "😊" in the comments section, which I can only interpret as overall satisfaction with the intervention. Their reports indicate that participants found significant value in the process, support, and structure of the intervention, even if all measured outcomes did not show substantial changes for each participant.

Contextualizing the Effect Size

Reviewing the Forest Plot (See Figure 11), each square represents a published intervention's standardized effect from the previous meta-analysis, with the bottom-most square depicting my SCED results. Specifically, the current study produced a small yet precise effect ($g = 0.08$; 95% CI [0.06, 0.10]). Compared to other studies, this intervention effect is generally lower in magnitude but exhibits a notably narrower confidence interval, with the narrow precision potentially reflecting the dense, within-participant measurement schedule inherent to ESM, through which I collected multiple data points per day to reduce variability and yield a more stable estimate.

Figure 11. Forest Plot w/Current Study



Methodological Considerations, Strengths, and Limitations

This study employed a rigorous multiple baseline across participants design, allowing for an idiographic analysis of intervention effects, revealing individual patterns of response that would be obscured in traditional group designs. The use of ESM for collecting primary outcome data (i.e., belonging, positivity) provided ecologically valid, real-time insights into participants' fluctuating states within their natural university environments. However, several limitations must be acknowledged.

The primary limitation is the small sample size (n = 2), which resulted from the early attrition of a third recruited individual, inherent to SCED, restricting the generalizability of findings to the broader population of autistic undergraduates. The observed effects, or lack thereof, are specific to Chase and Alyx. A secondary limitation is the aforementioned modifications to the PEERS®-YA intervention through the significant abbreviation of the intervention, removal of the social coaching aspect, and focus on one-to-one interventions rather than the intended group, resulting in a study that evaluates a substantially altered version of the established curriculum, further limiting conclusion about the efficacy of the PEERS®-YA program in this specific context. A tertiary limitation includes a relatively short maintenance phase for both participants, due to participant requests to finish the study before exams, which occurred immediately after each individual's intervention rather than as a longer, distinct post-semester phase. This, coupled with limited data points for Alyx's maintenance specifically, make it difficult to theorize firm conclusions about the long-term sustainment of any observed effects.

Furthermore, the reliance on self-report for all primary and secondary outcome measures, while providing valuable insight into internal states and perceptions, is subject to potential reporting biases. The significant variability in Alyx's daily data may reflect genuine fluctuations while simultaneously posing challenges for clear visual and statistical interpretation of intervention effects and highlighted the difficulty in establishing stable baselines within constrained timeframes. Finally, while visual analysis suggested some positive changes, specifically for Chase, the lack of statistical corroboration from the CDC method, potentially derived from its conservative innateness in connection to the observed data patterns, indicates an inconclusive result to the overall efficacy of the intervention under this specific context.

Implications for Practice and Future Research

Despite the inconclusive results regarding the effectiveness of this heavily modified intervention on ESM measures, this study offers several implications and highlights numerous avenues for future research.

Implications for Practice

Participants reported high social validity ratings, indicating that autistic university students appreciate the structured social learning opportunities and supportive program environments. Even if direct skill acquisition or ESM outcome changes were not uniformly significant, the perceived fairness, facilitator support, and suitability of the program were viewed as valuable outcomes. The significant influence of social partner presence, especially for Alyx, underscores the importance for practitioners to consider and leverage existing social contexts and supports when working with this population. These findings may suggest that highly abbreviated versions of comprehensive programs (e.g., PEERS®-YA) may not be sufficient in producing robust changes in complex outcomes such as daily belonging and positivity for all individuals, pointing to the need for adequate intervention dosage and further components (e.g., social coaches, group setting) to support generalization.

Implications for Future Research

The findings and limitations of my study emphasizes several important directions for future research aimed at enhancing social support for autistic university students. Specifically, replication and extension of this research with more diverse samples of autistic university students is warranted to improve the generalizability of findings, with future SCEDs benefiting from incorporating more participants or groups with multiple baseline designs. Beyond sample characteristics, intervention fidelity and dosage require further attention, with research investigating the PEERS®-YA curriculum implemented with full fidelity to its original model (i.e., all sessions, standard durations, social coaching) within traditional university settings to more accurately assess its potential, with

comparative studies examining different dosages or specific components that may yield valuable insights. The role of social context and confounding variables (i.e., social partner) merits more systematic investigation, with future studies measuring social context more directly as a covariate or exploring the efficacy of dyadic interventions. Building on the insights from the social validity responses and unexpected data patterns (i.e., dips, peaks) continued qualitative inquiry is essential to more deeply understand participants' lived experiences, individual differences in response to interventions, and their perspectives on impactful program elements.

Furthermore, the observed discrepancies between gains in social skills knowledge and changes in daily experiences, as seen with Alyx and Chase respectively, call for research into the complex relationship between social skills knowledge, demonstrated social skills performance (which was not measured in this study), and subjective daily social-emotional outcomes, potentially through naturalistic observational measures of social skills. Methodologically, the challenges experienced with certain statistical analyses in SCED in this study highlight the ongoing need for the continued exploration of method that effectively balance rigor with sensitivity to detect meaningful change, especially for outcomes that are inherently variable or exhibit gradual, subtle effects. To enhance robustness, future measurement approaches may benefit from developing or implementing outcome measures beyond self-report for variables like social engagement, belonging, or positivity through observational data, peer reports, or physiological indicators of stress or engagement. Finally, given the substantial heterogeneity in responses observed in this study, critical research into more individualized approaches, including but not limited to tailoring interventions to the specific baseline profiles, participant strengths, co-occurring conditions, and articulated goals of individual autistic undergraduates.

Chapter 5: Discussion and Future Research

For my dissertation, I embarked on a multi-faceted exploration of interventions designed to support autistic postsecondary students, with a particular focus on enhancing social engagement, fostering a sense of belonging, and understanding the efficacy of research methodologies exploring these domains. Through a sequential, three-study approach (i.e., systematic review, meta-analysis, SCED), I sought to build a comprehensive understanding, shifting from a broad synthesis of extant research to a focused empirical investigation. This chapter synthesizes the collective insights drawn from my efforts, discuss their integrated meaning, acknowledge the limitations of my work, and consider

Synthesized Discussion of Findings

My systematic literature review (i.e., Chapter 2) aimed to map the landscape of SCED research concerning interventions for autistic college students, given the growing prevalence and rate of enrollment observed in recent years. Through this review, I identified the types of interventions implemented and the behaviors they targeted (e.g., social, academic, adaptive), noting their general positive reported outcomes. This review also highlighted frequently perceived barriers to postsecondary success for autistic individuals, including but not limited to social engagement, academic integration, and independent living, examining how existing interventions implemented strategies to address these cited obstacles to inclusion. One of my key findings was methodological variability and limitations within the SCED literature, including inconsistencies in reporting procedural fidelity, assessing social validity, limited demographic reporting which restricts generalizability, and sometimes insufficient replication or detailed visual analysis procedures, highlighting the need for a quantitative synthesis to determine overall intervention effects and identify robust intervention characteristics.

Drawing from the results of the literature review, the meta-analysis (i.e., Chapter 3) provided a quantitative synthesis of SCED studies, with overall findings indicating a moderate-to-large and statistically significant benefit of interventions for autistic college students. However, this positive effect was accompanied by considerable heterogeneity, suggesting substantial variability in efficacy across studies, contexts, and intervention types. Further moderator analyses revealed identifying factors that significantly influenced intervention outcomes, with studies meeting NTACT:C Quality Indicators reporting more substantial gains, emphasizing the import of methodological rigor. Interventions implementing structured social planning and/or peer mediation and those utilizing technology-assisted supports (e.g., wearable technology), were similarly associated with larger effects. Additionally, interventions targeting social outcomes demonstrated consistent improvement. While the meta-analysis provided valuable insights into what interventions may work and for whom they may be effective, the meta-analysis had its own limitations, including but not limited to the persistent heterogeneity, evidence of potential publication bias, and removal of certain data, which were excluded for different reasons (i.e., baseline means equal to zero, fewer than three data points).

My final component in the SCED (i.e., Chapter 4) intended to offer an in-depth, ecologically valid evaluation of a heavily modified version of the PEERS®-YA curriculum, in efforts to enhance social engagement and examine the effects on sense of belonging and overall positivity for two autistic undergraduates in STEM or STEM-adjacent fields. Primary findings from ESM were largely inconclusive, specifically when analyzed through the statistically verifiable CDC method.

Despite the inconclusive outcomes from the ESM data, the pre-post measures offered a different perspective, with both participants reporting overall gains on the SCSR-M. One participant (i.e., Alyx) showed substantial gains in social skills knowledge via the TYASSK, whereas the other (i.e., Chase) showed minimal change but mentioned having prior skill development opportunities. Social validity ratings were notably high for both participants, indicating they valued the program's

fairness, facilitator support, and overall relevance. This SCED documented the complexities of implementing interventions in real-world university settings, the profound heterogeneity in autistic individuals' responses to intervention, and the nuanced challenges of measures subjective, internalized conditions (i.e., belonging, positivity) through intensive, longitudinal methodologies.

Importance of Null or Negative Findings

An emergent theme throughout my dissertation is the critical importance of reporting research that yields null, negative, or inconclusive findings. While studies with positive findings are the foundation of collective knowledge and facilitate effective practices, scientific methods are predicated on the principle of falsifiability, or the idea that for a hypothesis to be scientific, it must be possible to prove it wrong (Cook & Therrien, 2017; Therrien & Cook, 2018). However, a developing concern in special education is the presence of publication bias, which threatens the validity of the evidence base. Potential bias stems largely from the "file-drawer problem," where researchers opt not to submit studies with null results under the impression that journals hold bias against such findings, with an estimated 79% of school psychology faculty holding the belief that journals were less likely to publish null results and not worth submitting for publication (Iwachiw et al., 2019).

Researchers self-censoring creates a distorted evidence base with significant consequence, with underreporting of null, negative, or inconclusive findings making it more difficulty to identify the scope of an intervention's effectiveness (i.e., for which learners, under which contexts, and for which outcomes it may prove ineffective; Cook & Therrien, 2017; Walker & Carr, 2021). This practice may lead to an overestimation of intervention effects in research syntheses, potentially wasting finite resources on reexamining settled disputes regarding efficacy. Results from my meta-analysis investigation into potential publication bias aligns with this concern, which is especially acute for SCEDs which rely on replicable studies to establish statistical power. In a review of

SCEDs, 100% of reports exclusively documented positive findings (Therrien & Cook, 2018), underscoring the importance of transparent reporting for my SCED's largely inconclusive primary findings to address this issue. To combat the "file-drawer problem" and foster more robust science, the broader research community must continue to value methodologically sound studies regardless of the directions of their findings, a movement supported via journal policies encouraging the submission of null results and the adoption of "result neutral" review processes (Cook & Therrien, 2017; Therrien & Cook, 2018).

Limitations

While this dissertation aimed for methodological rigor across its three studies, several limitations must be acknowledged that temper the interpretation and generalizability of the findings. These limitations span the design and execution of the systematic review and meta-analysis, as well as the SCED study.

Systematic Literature Review and Meta-Analysis

Several methodological decisions within my systematic literature review and the subsequent meta-analysis warrant consideration. First, while IOA was established for the initial title/abstract screening and a portion of the final pool of included studies, I conducted aspects of the full-text screening independently. This approach introduces a potential risk, as research consistently indicates that single screening is not equivalent to double screening for study identification, with single versus double title/abstract screening resulting in a median of 5% missed studies (3% for experienced reviewers; Waffenschmidt et al., [2019]), which in some cases could substantially alter meta-analytic findings. More specific to my review, utilizing a second reviewer, even for experienced reviewers, at the full-text stage identified an additional 6.6% to 11.9% of eligible studies that the first reviewer had missed (Stoll et al., 2019). Given these findings, it is entirely possible that my independent review of some full texts may have led to inadvertent omission of potentially relevant studies.

Second, my review was limited to articles and dissertations exclusively published in English. Monolingual searches can significantly limit and bias the results of literature reviews, specifically for those aiming for broader, global relevance (Nuñez & Amano, 2021). There is a significant body of scientific literature that is published in languages other than English and ignoring these studies can result in skewed conclusions and the omission of important, context-dependent local evidence (Nuñez & Amano, 2021). My adoption of a monolingual search necessitates that I recognize and disclose that the exclusion of non-English literature may have led to the removal of significant international research, potentially affecting the overall comprehensiveness and applicability of the synthesized findings.

Single-Case Experimental Design

Designed to align with best practices in SCED, my study intended for a minimum of three replications to demonstrate experimental control (Horner et al., 2005; Kratochwill et al., 2010), which is a standard reflected in the NTACT:C criteria detailed in Chapter 2. However, due to participant attrition, my study proceeded with two participants. In this context, Lanovaz and Turgeon (2020) offer a critical perspective on the interpretation of findings employing a multiple baseline design. Questioning the empirical grounding of the strict three-demonstration rule, Lanovaz and Turgeon (2020) note the process could be overly stringent and lead to unacceptably low statistical power if an effect must be demonstrated across all tiers, with research indicating that, in a three-tier multiple baseline design, requiring a clear change in two or more tiers maintained adequate control of Type I errors while achieving acceptable power. While pertaining to designs with at least three tiers, the broader argument of demanding perfect replication across all planned tiers may increase the burden of proof, resulting in the dismissal of true, albeit smaller or variable, effects and encourage researchers to consider limitations to power when requiring all tiers show a clear, distinctive change (Lanovaz & Turgeon, 2020).

Apart from acknowledgements to internal validity and statistical power within the study, the contribution of my SCED (n = 2) can be measured through the perspective of how generalizability is achieved in single-case research. Walker and Carr (2021) argue against the common misconception that generalizations from SCEDs are primarily limited to the number of participants within that specific study, positing that single-case research facilitates generalizations through a line of replication across multiple studies rather than a single, large-*N* study. This replication framework suggests that knowledge and generalizability accrue through direct, systematic replications across multiple studies, collectively forming a line of replicable SCEDs (Walker & Carr, 2021). While my SCED with two participants inherently had fewer opportunities to demonstrate direct replication of effects within the study, my procedures and findings, including but not limited to the detailed descriptions of participant demographics, specific modifications to the PEERS®-YA curriculum, context of a midwestern university serving multiple communities, and the nuanced individual outcomes, contribute valuable data to the broader line of research on interventions for autistic university students. By providing detailed reporting, I allow future researchers to assess the similarity of their own clients and circumstances to those in my study, informing the iterative process through which the broader field of single-case research establishes the generality of effects.

Future Research

The collective findings from this dissertation underscore the multifaceted nature of the autistic student experience in higher education and illuminate several critical avenues for future inquiry. While my research focused on intervention efficacy and social engagement, the data revealed the profound influence of adjacent domains that warrant more dedicated investigation. Specifically, future research should prioritize a deeper exploration of identity formation among autistic university students, the crucial role of targeted mental health supports in fostering well-being and persistence, and the intricate connection between executive functioning and social engagement. Investigating

these areas is essential for developing the holistic, person-centered support systems that enable autistic students to not only succeed academically but to thrive personally and socially within the university environment and into adulthood.

Identity Formation

Conducting my dissertation, particularly the intensive work with individual participants and collaborative efforts with my colleagues, instilled in me a new line of inquiry that is critical to the field, specifically in adequately and appropriately serving autistic postsecondary students. My experiences highlighted that, to sufficiently support autistic students, the field must look beyond a singular concept of identity and consider the complex nature of identity formation, especially during the transition to university life. My initial conceptual exploration into this domain, which forms the basis for this proposed research, began by examining the intersection of autistic identity with other specific marginalized identities. That preliminary review of research, which synthesized research on how autistic individuals navigate multidimensional aspects of self, emphasized the broader need for research into the dynamics of developing an authentic identity within the complex social environment of a university.

Therefore, in future research, I intend to use my preliminary research as a foundation to investigate this intricate process of identity formation for autistic students, including but not limited to the tension between developing an authentic self-concept and the pressure to mask or conform to neurotypical expectations. For future studies, I plan to examine how stigmatizing social environments can disrupt this pursuit of authenticity, leading to internalized stigma and isolation, Conversely, I want to explore how access to affirming, autistic-led peer and research communities can foster resilience and positive mental health outcomes by providing spaces where students feel empowered to simply be themselves, informing the creation of a truly inclusive college experience.

Mental Health

Through my extensive work with individual participants, it became evident that there's a significant connection between a student's campus experiences and their mental well-being, pointing to a critical need for further investigation. While my research focused on social engagement, it became clear that the emotional and psychological health of autistic students, particularly in relation to negative social encounters, is a foundational element of their ability to persist and thrive in postsecondary education.

For future research, I intend to explore the development and implementation of holistic mental health supports tailored for autistic individuals in postsecondary settings, including traditional universities, community colleges, and inclusive postsecondary education programs. During this inquiry, I should recognize that mental health challenges for this population are not simply internal or diagnostic issues, perhaps exacerbated by the immense social and academic pressures of the university environment, with the stress of navigating complex social landscapes and the perceived need to mask autistic traits taking a significant toll on overall well-being. Consequently, I want to evaluate support systems that extend beyond traditional counseling services, including but not limited to preventative, campus-wide initiatives that examine **(a)** how to create more inclusive and predictable social environments and **(b)** how to train faculty and staff to be effective allies, potentially reducing the external stressors that contribute to poor mental health.

Executive Functioning

During my study, I conducted preliminary research to examine executive functioning as a critical variable influencing student outcomes, justifying a closer investigation into its foundational role in social success. Many of the complex demands of social life on a college campus (e.g., reciprocal conversation management, shifting conversation topics, planning group activities, regulating emotional responses) rely heavily on underlying executive functioning skills, including

cognitive flexibility, planning, and self-monitoring. Without a strong foundation in these crucial skills, autistic students may find the fluid and unpredictable dynamics of social engagement particularly demanding, potentially hindering their ability to form meaningful connections and foster a true sense of belonging. I would like to explore interventions that focus less on discrete social scripts in favor of strengthening core executive functioning capacities that underpin social competencies, with this line of inquiry seeking to determine if enhancing these foundational cognitive skills can lead to more flexible, generalized, and authentic social engagement for autistic students in their daily lives.

References

*Denotes the document was included in the systematic review.

†Denotes the document was included in the meta-analysis.

Adreon, D. & Durocher, J. S. (2007). Evaluating the college transition needs of individuals with high-functioning autism spectrum disorders. *Intervention in School and Clinic*, *42*(5), 271-279. https://doi.org/10.1177/10534512070420050201

Alkan, N. (2016). Psychological Sense of University Membership: An adaptation study of the PSSM scale for Turkish university students. *The Journal of Psychology, 150*(4), 431-499. https://doi.org/10.1080/00223980.2015.1087373

Anderson, A. H., Stephenson, J., & Carter, M. (2017). A systematic literature review of the experiences and supports of students with autism spectrum disorder in post-secondary education. *Research in Autism Spectrum Disorders*, *39*, 33-53. http://dx.doi.org/10.1016/j.rasd.2017.04.002

*†Ashbaugh, K., Koegel, R., & Koegel, L. (2017). Increasing social integration for college students with autism spectrum disorder. *Behavioral Development Bulletin, 22*(1), 183–196. https://doi.org/10.1037/bdb0000057

Atz, U. (2013). Evaluating experience sampling of stress in a single-subject research design. *Personal and Ubiquitous Computing, 17*(4), 639-652. https://ssrn.com/abstract=3476700

Bakker, T., Krabbendam, L., Bhulai, S., Meeter, M., & Begeer, S. (2023). Study progression and degree completion of autistic students in higher education: a longitudinal study. *Higher Education, 85*(1), 1-26. https://doi.org/10.1007/s10734-021-00809-1

Bentley, K. H., Maimone, J. S., Kilbury, E. N., Tate, M. S., Wisniewski, H., Levine, M. T., Roberg, R., Torous, J. B., Nock, M. K., & Kleiman, E. M. (2021). Practices for monitoring and responding to incoming data on self-injurious thoughts and behaviors in intensive

longitudinal studies: A systematic review. *Clinical Psychology Review, 90*, 102098. https://doi.org/10.1016/j.cpr.2021.102098

Black, M. H., Milbourn, B., Chen, N. T. M., McGarry, S., Wali, F., Ho, A. S. V., Lee, M., Bölte, S., Falkmer, T., & Girdler, S. (2020). The use of wearable technology to measure and support abilities, disabilities, and functional skills in autistic youth: A scoping review. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology, 8*(1), 48-69. https://doi.org/10.21307/sjcapp-2020-006

Bulté, I., & Onghena, P. (2009). Randomization tests for multiple-baseline designs: An extension of the SCRT-R package. *Behavior Research Methods, 41*(2), 477-485. https://doi.org/10.3758/brm.41.2.477

Carter, S. L. (2009). *The social validity manual: A guide to subjective evaluation of behavior interventions*. Elsevier, Inc. https://doi.org/10.1016/C2017-0-03852-1

Cheriyan, C., Shevchuk-Hill, S., Riccio, A., Vincent, J., Kapp, S. K., Cage, E., Dwyer, P., Kofner, B., Attwood, H., & Gillespie-Lynch, K. (2021). Exploring the career motivations, strengths, and challenges of autistic and non-autistic university students: Insights from a participatory study. *Frontiers in Psychology*, 12:719827. https://doi.org/10.3389/fpsyg.2021.719827

Chien, Y.-L., Tsai, W.-C., Chen, W.-H., Yang, C.-L., Shur-Fen Gau, S., Soong, W.-T., Laugeson, E. A., & Chiu, Y.-N. (2021). Effectiveness, durability, and clinical correlates of the PEERS® social skills intervention in young adults with autism spectrum disorder: The first evidence outside North America. *Psychological Medicine, 53*, 966-976. https://doi.org/10.1017/S0033291721002385

Cobian, K. P., White-Lewis, D. K., Hurtado, S., & Ramos, H. V. (2022). Implementing case study design to evaluate diverse institutions and STEM education contexts: Lessons and key areas for systematic study. *New Directions for Evaluation, 174*, 21-31.

https://doi.org/10.1002/ev.20507

Cook, B. G., & Therrien, W. J. (2017). Null effects and publication bias in special education research. *Behavioral Disorders, 42*(4), 149-158. https://doi.org/10.1177/0198742917709473

Crabtree, A., Caudel, D., Pinette, J., Vang, C., Neikirk, K., Kabugi, K., Zaganjor, E., & Hinton, A. (2024). Recruiting and retaining autistic talent in STEMM. *iScience, 27*(3), 109080. https://doi.org/10.1016/j.isci.2024.109080

Daily, S. B. & Eugene, W. (2013). Preparing the future STEM workforce for diverse environments. *Urban Education, 48*(5), 682-704. https://doi.org/10.1177/0042085913490554

*†Detar, W. J. & Vernon, T. W. (2020) Targeting question-asking initiations in college students with ASD using a video-feedback intervention. *Focus on Autism and Other Developmental Disabilities, 35*(4), 208-220. https://doi.org/10.1177/1088357620943506

Dietz, P.M., Rose, C.E., McArthur, D., & Maenner, M. (2020). National and state estimates of adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 50*(12), 4258–4266. https://doi.org/10.1007/s10803-020-04494-4

Doggett, R., Krasno, A. M., Koegel, L., & Koegel, R. L. (2013). Acquisition of multiple questions in the context of social conversation in children with autism. *Journal of Autism and Developmental Disorders, 43*(9), 2015-2025. https://doi.org/10.1007/s10803-012-1749-8

Duerksen, K., Besney, R., Ames, M. E., & McMorris, C. A. (2021). Supporting autistic adults in postsecondary settings: A systematic review of peer mentorship programs. *Autism in Adulthood, 3*(1), 85-99. https://doi.org/10.1089/aut.2020.0054

Dwyer, P., Mineo, E., Mifsud, K., Lindholm, C., Gurba, A., & Waisman, T. C. (2023). Building neurodiversity-inclusive postsecondary campuses: Recommendations for leaders in higher education. *Autism in Adulthood*, *5*(1), 1-14. https://doi.org/10.1089/aut.2021.0042

*†Engel, S. E. (2011). *An animal-assisted intervention with college students with Asperger's Syndrome* (Publication No. 3444111) [Doctoral dissertation, Alfred University]. ProQuest Dissertations and Theses Global. http://hdl.handle.net/10829/2668

*†Engstrom, E. T. (2019). *Improving daily living skills in college students with autism spectrum disorder using a peer-mediated daily living checklist intervention* (Publication No. 22592311) [Doctoral dissertation, University of California - Santa Barbara]. ProQuest Dissertations and Theses Global. https://escholarship.org/uc/item/6h81r4th

Fisher, W. W., Kelley, M. E., & Lomas, J. E. (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis, 36*(3), 387-406. https://doi.org/10.1901/jaba.2003.36-387

Fritz, J., Piccirillo, M. L., Cohen, Z. D., Frumkin, M., Kirtley, O., Moeller, J., Neubauer, A. B., Norris, L. A., Schuurman, K. N., Snippe, E., & Bringmann, L. F. (2024). So you want to do ESM? 10 essential topics for implementing the experience-sampling method. *Advances in Methods and Practices in Psychological Science, 7*(3). https://doi.org/10.1177/25152459241267912

Fullagar, C. J. & Kelloway, E. K. (2009). 'Flow' at work: An experience sampling approach. *Journal of Occupational and Organizational Psychology, 82*, 594-615. https://doi.org/10.1348/096317908X357903

Fung, L. (Ed.). (2021). *Neurodiversity: From phenomenology to neurobiology and enhancing technologies*. American Psychiatric Association Publishing.

*Gunn, S., Sellers, T. P., & Lignugaris, Kraft, B. (2017). Application of coaching and behavioral skills training during a preschool practicum with a college student with autism spectrum disorder. *Clinical Case Studies, 16*(4), 275-294. https://doi.org/10.1177/1534650117692673

Hamilton, L. G. & Petty, S. (2023). Compassionate pedagogy for neurodiversity in higher education: A conceptual analysis. *Frontiers in Psychology*, 14:1093290.

https://doi.org/10.3389/fpsyg.2023.1093290

Harvery, M., Froude, E., Foley, K., Trollor, J. N., & Arnold, S. R. C. (2021). Employment profiles of autistic adults in Australia. *Autism Research, 14*(10), 2061-2077. https://doi.org/10.1002/aur.2588

Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2012). A standardized mean difference effect size for single-case designs. *Research Synthesis Methods, 3*(3), 224–239. https://doi.org/10.1002/jrsm.1052

Honan, I., Sharp, N., McIntyre, S., Smithers-Sheedy, H., Balde, I., Quinn, K., Morgan, M., Rothery, S., Butchers, T., & Laugeson, E. A. (2023). Program evaluation of an adapted PEERS® social skills program in young adults with autism spectrum disorder and/or mild intellectual impairments and social skills difficulties. *Journal of Evaluation in Clinical Practice, 29*(1), 126-135. https://doi.org/10.1111/jep.13743

Hott, B. L., Brigham, F. J., & Peltier, C. (2021). *Research methods in special education*. SLACK Incorporated.

Howorth, S. K., Rooks-Ellis, D. L., Cobo-Lewis, A., Taylor, J., & Moody, C. T. (2022). Effects of an abbreviated and adapted PEERS® curriculum as part of a college transition program for young adults on the autism spectrum. *Career Development and Transition for Exceptional Individuals, 46*(2), 69-82. https://doi.org/10.1177/21651434221098411

Howorth, S. K., Rooks-Ellis, D. L., Sulinski, E., & Jones, B. (2023). Effect of PEERS® class on conversational skills of adults with autism in college. *Journal of Postsecondary Education and Disability, 35*(4), 273-288.

Hua, Y., Zhao, Q., Shen, J., Liu, Y., Zheng, L., & Zhang, M. (2022). Program for education and enrichment of relational skills (PEERS) training for social skills and depressed mood intervention in young adult with depression: Study protocol for a randomized controlled

trial. *Frontiers in Psychiatry, 13*. https://doi.org/10.3389/fpsyt.2022.993124

Institution of Education Sciences (n.d.). *Postsecondary Education*. https://ies.ed.gov/postsecondary-education

Iwachiw, J. S., Button, A. L., & Atlas, J. (2019). The perceived role of null results in school psychology research and publication. *School Psychology, 40*(4), 416-430. https://doi.org/10.1177/0143034319851230

*†Jackson, L. G., Duffy, M. L., Brady, M. P., & McCormick, J. (2018). Effects of learning strategy training on the writing performance of college students with Asperger's syndrome. *Journal of Autism and Developmental Disorders, 48*(3), 708-721. https://doi.org/10.1007/s10803-017-3170-9

*†Koegel, L. K., Ashbaugh, K., Koegel, R. L., & Detar, W. J. (2013). Increasing socialization in adults with Asperger's Syndrome. *Psychology in the Schools, 50*(9), 899-909. https://doi.org/10.1002/pits.21715

*†Koegel, L. K., Navab, A., Ashbaugh, K., & Koegel, R. L. (2015). Using reframing to reduce negative statements in social conversation for adults with autism spectrum disorder. *Journal of Positive Behavior Interventions, 18*(3), 133-144. https://doi.org/10.1177/1098300715596136

*†Koegel, L. K., Ashbaugh, K., Navab, A., & Koegel, R. L. (2016). Improving empathic communication skills in adults with autism spectrum disorder. *Journal on Autism and Developmental Disorders, 46*(3), 921-933. https://doi.org/10.1007/s10803-015-2633-0

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M. & Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse. http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Kratochwill, T. R., Horner, R. H., Leving, J. R., Machalicek, W., Ferron, J., & Johnson, A. (2023). Single-case intervention research design standards: Additional proposed upgrades and future

directions. *Journal of School Psychology, 97*, 192-216. https://doi.org/10.1016/j.jsp.2022.12.002

Lanovaz, M. J. & Turgeon, S. (2020). How many tiers do we need? Type I errors and power in multiple baseline designs. *Perspectives on Behavior Science, 43*(3), 605-616. https://doi.org/10.1007/s40614-020-00263-x

Laugeson E. A. (2017). *PEERS® for young adults: social skills training for adults with autism spectrum disorder and other social challenges*. Taylor and Francis. https://doi.org/10.4324/9781315297057

Laugeson, E. A., Gantman, A., Kapp, S. K., Orenski, K., & Ellingsen, R. (2015). A randomized controlled trial to improve social skills in young adults with autism spectrum disorder: The UCLA PEERS® Program. *Journal of Autism and Developmental Disorders, 45*(12), 3978-3989. https://doi.org/10.1007/s10803-015-2504-8

Ledford, J. R., Lane, J. D., & Severini, K. E. (2017). Systematic use of visual analysis for assessing outcomes in single case design studies. *Brain Impairment, 19*(1), 4-17. https://doi.org/10.1017/BrImp.2017.16

Lei, J., & Russell, A. (2021). Understanding the role of self-determination in shaping university experiences for autistic and typically developing students in the United Kingdom. *Autism, 25*(5), 1262-1278. https://doi.org/10.1177/1362361320984897

Lisnyj, K. T., Gillani, N., Pearl, D. L., McWhirter, J. E., & Papadopoulos, A. (2023). Factors associated with stress impacting academic success among post-secondary students: A systematic review. *Journal of American College Health, 71*(3), 851–861. https://doi.org/10.1080/07448481.2021.1909037

MacDonald, K. (2018). A review of the literature: The needs of nontraditional students in postsecondary education. *Strategic Enrollment Management Quarterly, 5*(4), 159–164. https://doi.org/10.1002/sem3.20115

Madaus, J., Gelbar, N., Faggella-Luby, M. N., & Dukes III, L. L. (2024). Ten years after: A

systematic review of the literature on postsecondary education and disability from 2013 to 2022. *Frontiers in Education,* 9:1412903. https://doi.org/10.3389/feduc.2024.1412903

Maenner, M. J., Warren, Z., Robinson Williams, A., Amoakohene, E., Bakian, A. V., Bilder, D. A., Durkin, M. S., Fitzgerald, R. T., Furnier, S. M., Hughes, M. M., Ladd-Acosts, C. M., McArthur, D., Pas, E. T., Salinas, A., Vehorn, A., Williams, S. Esler, A., Grzybowski, A., Hall-Lande, J., . . . Shaw, K. A. (2023). Prevalence and characteristics of Autism Spectrum Disorder among children aged 8 years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2020. *MMWR Surveillance Summaries, 72*(2), 1–14. http://dx.doi.org/10.15585/mmwr.ss7202a1

*†Mann, C. C. & Karsten, A. M. (2020). Efficacy and social validity of procedures for improving conversational skills of college students with autism. *Journal of Applied Behavioral Analysis, 53*(1), 402-421. https://doi.org/10.1002/jaba.600

*†Mason, R. A., Rispoli, M., Ganz, J. B., Boles, M. B., & Orr, K. (2012). Effects of video modeling on communicative social skills of college students with Asperger Syndrome. *Developmental Neurorehabilitation, 15*(6), 425-434. https://doi.org/10.3109/17518423.2012.704530

Mazereel, V., Vansteelandt, K., Menne-Lothmann, C., Decoster, J., Derom, C., Thiery, E., Rutten, B. P. F., Jacobs, N., van Os, J., Wichers, M., De Hert, M., Vancampfort, D., & van Winkel, R. (2021). The complex and dynamic interplay between self-esteem, belongingness, and physical activity in daily life: An experience sampling study in adolescence and young adulthood. *Mental Health and Physical Activity, 21*, 100413. https://doi.org/10.1016/j.mhpa.2021.100413

*†McMahon, D. D., Cihak, D. F., Wright, R. E., Bell, & Bell, S. M. (2016). Augmented reality for teaching science vocabulary to postsecondary education students with intellectual disabilities and autism. *Journal of Research on Technology in Education, 48*(1), 38-56.

https://doi.org/10.1080/15391523.2015.1103149

McVey, A. J., Dolan, B. K., Willar, K. S., Pleiss, S., Karst, J. S., Casnar, C. L., Caiozzo, C., Vogt, E. M., Gordon, N. S., & Vaughan Van Hecke, A. (2016). A replication and extension of the PEERS® for Young Adults social skills intervention: Examining effects on social skills and social anxiety in young adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 46*(12), 3739-3754. https://doi.org/10.1007/s10803-016-2911-5

McVey, A. J., Schiltz, H., Haendel, A., Dolan, B. K., Willar, K. S., Pleiss, S., Karst, J. S., Carson, A. M., Caiozzo, C., Vogt, E., & Vaughan Van Hecke, A. (2017). Brief report: Does gender matter in intervention for ASD? Examining the impact of the PEERS® social skills intervention on social behavior among females with ASD. *Journal of Autism and Developmental Disorders, 47*(7), 2282-2289. https://doi.org/10.1007/s10803-017-3121-5

Mölsä, M. E., Lax, M., Korhonen, J., Gumpel, T. P., and Söderberg, P. (2022). The experience sampling method in monitoring social interactions among children and adolescents in school: A systematic literature review. *Frontiers in Psychology, 13*, 844698. https://doi.org/10.3389/fpsyg.2022.844698

Morningstar, M. E., Bassett, D. S., Cashman, J., Kochhar-Bryant, C., & Wehmeyer, M. L. (2012). Aligning transition services with secondary educational reform: A position statement of the Division on Career Development and Transition. *Career Development and Transition for Exceptional Individuals, 35*(3), 132-142. https://doi.org/10.1177/2165143412454915

Nail, E. (2024). Factors in successful programme completion for postsecondary students with disabilities: A systematic review of the literature. *British Journal of Special Education, 51*(3), 346-357. https://doi.org/10.1111/1467-8578.12539

National Center for Education Statistics. (2022). Undergraduate Retention and Graduation Rates.

Condition of Education. U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/programs/coe/indicator/ctr

Nelson, H., Switalsky, D., Ciesielski, J., Brown, H. M., Ryan, J., Stothers, M., Coombs, E., Crerear, A., Devlin, C., Bendevis, C., Ksiazek, T., Dwyer, P., Hack, C., Connolly, T., Nicholas, D. B., & DiRezze, B. (2023). A scoping review of supports on college campuses for autistic post-secondary students. *Frontiers in Education,* 8:1179865. https://doi.org/10.3389/feduc.2023.1179865

Newman, L., Wagner, M., Knokey, A. M., Marder, C., Nagle, K. Shaver, D., Wei, X., Cameto, R., Contrera, E., Ferguson, K., Greene, S., & Schwarting, M. (2011). *The post-high school outcomes of young adults with disabilities up to 8 years after high school: A report from the National Longitudinal Transition Study-2* (NLTS2) (NCSER 2011-3005). Institute on Education Sciences. https://files.eric.ed.gov/fulltext/ED524044.pdf

Nicholas, D. B., Attridge, M., Zwaigenbaum, L., & Clarke, M. (2014). Vocational support approaches in autism spectrum disorder: a synthesis review of the literature. *Autism, 19*(2), 235-245. https://doi.org/10.1177/1362361313516548

Nuñez, M. A., & Amano, T. (2021). Monolingual searches can limit and bias results in global literature reviews. *Nature Ecology & Evolution, 5*(3), 264. https://doi.org/10.1038/s41559-020-01369-w

Oh, M., Laugeson, E. A., Kim, J.-H., Lee, K., Kim, J., Lee, S., Lim, B., Cha, S., Bong, G., Yoon, N., Bahn, G. H., & Yoo, H. J. (2021). A randomized controlled trial of the Korean version of the Program for the Education and Enrichment of Relational Skills for Young Adults (PEERS®-YA-K) with autism spectrum disorder: A pilot study. *Frontiers in Psychiatry, 12*. https://doi.org/10.3389/fpsyt.2021.730448

Ohl, A., Sheff, M. G., Small, S., Nguyen, J., Paskor, K., & Zanjirian, A. (2017). Predictors of

employment status among adults with autism spectrum disorder. *Work, 56*(2), 345-355. https://doi.org/10.3233/wor-172492

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A, Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … & McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ, 372*, n160. https://doi.org/10.1136/bmj.n160

Parker, R. I. & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40*(4), 357-367. https://doi.org/10.1016/j.beth.2008.10.006

Petcu, S. D., Zhang, D., & Li, Y. F. (2021). Students with autism spectrum disorders and their first-year college experiences. *International Journal of Environmental Research and Public Health*, *18*(22), 11822. https://doi.org/10.3390/ijerph182211822

Platos, M., Wojaczek, K., & Laugeson, E. A. (2023). Fostering friendship and dating skills among adults on the autism spectrum: A randomized controlled trial of the Polish version of the PEERS® for Young Adults curriculum. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-023-05921-y

Pustejovsky, J. E. (2018). Using response ratios for meta-analyzing single-case designs with behavioral outcomes. *Journal of School Psychology, 68*, 99-112. https://doi.org/10.1016/j.jsp.2018.02.003

Pustejovsky, J. E., Chen, M., Grekov, P., & Swan, D. M. (2024). Single-case effect size calculator (Version 0.7.3) [Web Application]. https://jepusto.shinyapps.io/SCD-effect-sizes/

Pustejovsky, J. E., Chen, M., Hamilton, B., & Grekov, P. (2023). scdhlm: A web-based calculator for between-case standardized mean differences (Version 0.7.2) [Web Application].

https://jepusto.shinyapps.io/scdhlm

Pustejovsky, J. E., Hedges, L. V., & Shadish, W. R. (2014). Design-comparable effect sizes in multiple baseline designs: A general modeling framework. *Journal of Educational and Behavioral Statistics, 39*(5), 368-393. https://doi.org/10.3102/1076998614547577

Raymaker, D., Sharer, M., Maslak, J., Powers, L. E., McDonald, K., Kapp, S. K., Moura, I., Wallington, F., & Nicolaidis, C. (2022). "[I] don't wanna just be like a cog in the machine": Narratives of autism and skilled employment. *Autism, 27*(1), 65-75. https://doi.org/10.1177/13623613221080813

*Reed, D. K., Hallett, A., & Rimel, H. (2016). Note-taking instruction for college students with autism spectrum disorder. *Exceptionality, 24*(4), 195-212. http://dx.doi.org/10.1080/09362835.2015.1107833

Renbarger, R., & Long, K. (2019). Interventions for postsecondary success for low-income and high-potential students: A systematic review. *Journal of Advanced Academics, 30*(2), 178–202. https://doi.org/10.1177/1932202X19828744

Rizvi, S. L., & Nock, M. K. (2008). Single-case experimental designs for the evaluation of treatments for self-injurious and suicidal behaviors. *Suicide and Life-Threatening Behavior, 38*(5), 498-510. https://doi.org/10.1521/suli.2008.38.5.498

Rose, A. J., Kelley, K. R., & Raxter, A. (2021). Effects of PEERS® social skills training on young adults with intellectual and developmental disabilities. *Behavior Modification, 45*(2), 297-323. https://doi.org/10.1177/0145445520987146

Roux, A. M., Koffer-Miller, K. H., Tao, S., Rast, J. E., Ventimiglia, J., Shattuck, P., & Shea, L. (2023). Unrealized cross-system opportunities to improve employment and employment-related services among autistic individuals. *The Milbank Quarterly, 101*(4), 1223-1279. https://doi.org/10.1111/1468-0009.12666

Sabitelli, R. M., & Anderson, S. A. (2005). *Assessing outcomes in child and youth programs: A practical handbook—revised edition*. University of Connecticut, Center for Applied Research.

Sanson-Fisher, R., D'Este, C., Carey, M., Noble, N., & Paul, C. (2014). Evaluation of systems-oriented public health interventions: alternative research designs. *Annual Review of Public Health, 35*(1), 9-27. https://doi.org/10.1146/annurev-publhealth-032013-182445

Schembri-Mutch, T. M., McCrimmon, A. W., & Zwiers, M. L. (2023). Understanding the needs of post-secondary students. *Journal of Autism and Developmental Disabilities, 54*(7), 2636-2650. https://doi.org/10.1007/s10803-023-06010-w

Scott, M., Milbourn, B., Falkmer, M., Black, M. H., Bölte, S., Halladay, A., Lerner, M. D., Taylor, J. L., & Girdler, S. (2018). Factors impacting employment for people with autism spectrum disorder: A scoping review. *Autism, 23*(4), 869-901. https://doi.org/10.1177/1362361318787789

*†Siko, L. (2018). *Wearable technology for self-monitoring of academic engagement in college students with autism spectrum disorder* (Publication No. 10814889) [Doctoral Dissertation, George Mason University. ProQuest Dissertations and Theses Global. https://eric.ed.gov/?id=ED587356

Stewart, K. K., Carr, J. E., Brandt, C. W., & McHenry, M. M. (2007). An evaluation of the conservative dual-criterion method for teaching university students to visually inspect AB-design graphs. *Journal of Applied Behavior Analysis, 40*(4), 713-718. https://doi.org/10.1901/jaba.2007.713–718

Stoll, C., Izadi, S., Fowler, S., Green, P., Suls, J., & Colditz, G. A. (2019). The value of a second reviewer for study selection in systematic reviews. *Research Synthesis Methods, 10*(4), 539–545. https://doi.org/10.1002/jrsm.1369

Swoboda, C. M., Kratochwill, T. R., & Levin, J. R. (2010). *Conservative dual-criterion method for single-case research: A guide for visual analysis of AB, ABAB, and multiple-baseline designs* (WCER Working

Paper No. 2010-13). University of Wisconsin–Madison, Wisconsin Center for Education Research. https://wcer.wisc.edu/docs/working-papers/Working_Paper_No_2010_13.pdf

Taylor, J. L., Henninger, N. A., & Mailick, M. R. (2015). Longitudinal patterns of employment and postsecondary education for adults with autism and average-range IQ. *Autism, 19*(7), 785-793. https://doi.org/10.1177/1362361315585643

Therrien, W. J., & Cook, B. G. (2018). Introduction to special issue: Null effects and publication bias in learning disabilities research. *Learning Disabilities: Research & Practice, 33*(1), 5-10. https://doi.org/10.1111/ldrp.12163

Tomaszewski, B., Klinger, L., & Pugliese, C. E. (2022). Self-determination in autistic transition-aged youth without intellectual disability. *Journalism of Autism and Developmental Disorders, 52*(9), 4067-4078. https://doi.org/10.1007/s10803-021-05280-6

Verhagen, S. J. W., Hasmi, L., Drukker, M., van Os, J., & Delespaul, P. A. E., G. (2016). Use of the experience sampling method in the context of clinical trials. *Evidence-Based Mental Health, 19*(3), 86-89. https://doi.org/10.1136/ebmental-2016-102418

Valentine, J. C., Tanner-Smith, E. E., Pustejovsky, J. E., & Lau, T. S. (2016). Between-case standardized mean difference effect sizes for single-case designs: A primer and tutorial using the scdhlm web application. *Campbell Systematic Reviews, 12*(1), 1–31. https://doi.org/10.4073/cmdp.2016.1

von Klipstein, L., Servaas, M. N., Schoevers, R. A., van der Veen, D. C., & Riese, H. (2023). Integrating personalized experience sampling in psychotherapy: A case illustration of the Therap-i module. *Heliyon, 9*(3), e14507. https://doi.org/10.1016/j.heliyon.2023.e14507

Waffenschmidt, S., Knelangen, M., Sieben, W., Bühn, S., & Pieper, D. (2019). Single screening versus conventional double screening for study selection in systematic reviews: a methodological systematic review. *BMC Medical Research Methodology, 19*(1), 132.

https://doi.org/10.1186/s12874-019-0782-0

Willson, R. J., Veale, D., & Freeston, M. H. (2016). Imagery rescripting for body dysmorphic disorder: A multiple-baseline single-case experimental design. *Behavior Therapy, 47*(2), 248-261. https://doi.org/10.1016/j.beth.2015.08.006

*†Wright, R. E., McMahon, D. D., Cihak, D. F., &. Hirschfelder, K. (2022). Smartwatch executive function supports for students with ID and ASD. *Journal of Special Education Technology, 37*(1), 63-73. https://doi.org/10.1177/0162643420950027

Yassa, M. O. & Chen, C. (2023). Outcomes and lessons learned from an intensive summer undergraduate research program in neuroscience. *The Journal of STEM Outreach, 6*(1). https://doi.org/10.15695/jstem/v6i1.06

Appendix A: Literature Review and Meta-Analysis (Tables and Figures)

Table 1. Research Design

Author (Year)	Design	Setting	Independent Variable	Dependent Variable(s)
Ashbaugh et al. (2017)	MBDP	University; Community College	Structured social planning: (a) motivational interests (b) menu of social activities (c) organizational skills (d) peer mentors (e) social skills training	On-campus social integration: (a) number of college and community-based social activities attended per week (b) number of extracurricular activities attended per week (c) cumulative number of different peers interacted with at social activities *(d) supplemental measure of academic performance* *(e) social validation*
Detar (2020)	MBDP	University; Community College	Social inquiry: (a) video-feedback (b) practice conversation	Social behaviors: (a) questions asked (b) perseverative discussion (c) conversational pauses (d) affect/interest (e) minimal conversational reciprocity
Engel (2009)[a]	MBDP	University	Animal-Assisted Therapy	(a) attention toward consultant (b) attention toward therapy dog (c) appropriate verbalizations (d) repetitive behaviors (e) session length
Engstrom (2019)[a]	MBDP	University; Community College	Checklist intervention with peer mediation	Percentage of daily living skills completed: (a) household maintenance

				(b) money-management (c) organization (d) self-care (e) use of time (f) personal hygiene *Adaptive behavior, mental health, and quality of life:* *(a) Vineland Adaptive Behavior Scales-3rd Edition* *(b) Beck Depression Inventory-2nd Edition* *(c) Beck Anxiety Inventory* *(d) Quality of Life Assessment for Adults with ASD* (e) *Quarterly Grade Report*
Gunn et al. (2017)	ABCD	University	Behavior skills training: (a) social pragmatics (b) executive functioning	Classroom management: (a) visual scanning (b) verbal interactions (c) student engagement
Jackson et al. (2018)	MBDP	University	Writing intervention (DATE + Self-Regulated Strategy Development [SRSD]): • **D**evelop thesis statement • **A**dd supporting details • **T**ie it together • **E**dit and revise with SCOPES o **S**entences o **C**apitalization o **O**rder o **P**unctuation o **E**dit o **S**pelling	Quality of the written product: (a) opening (thesis statement) (b) argument (supporting details) (c) organization and structure (d) conclusion (e) disciplinary concerns (format and citations) (f) grammar and syntax Secondary behavioral measures • *Total number of correct words written* • *Duration of time spent writing*

Koegel et al. (2013)	MBDP	University	Structured social planning: (a) assessment of social activities (b) researching community and/or university events (c) organizational skills (d) peer mentors (e) weekly meetings (f) systematic fading of support	Social activities: (a) activity log (b) peer mentor verification
Koegel et al. (2016a)	MBDP	University	Reframing of negative statements: (a) role-play scenarios (b) video recording and video-feedback sessions (c) videotaped conversation a. self-monitoring	Social conversation: (a) negative statements (b) positive/neutral statements (c) affect and interest during conversation (d) social desirability
Koegel et al. (2016b)	MBDP	University	Conversation partners: (a) video-feedback (b) visual framework	Empathic responses: (a) percentage of responses including an empathic statement (b) percentage of responses including an empathic question (c) collateral gains in levels of empathy (d) confidence in peer interactions
Mann & Karsten (2020)	MBDB	University	Conversational skills: (a) problem detection (b) solution provision (c) self-questioning training (d) peer-led probes (e) natural environment tests	Correct responses: • based on observations of neurotypical peers to gauge targeted skills

Mason et al. (2012)	MBDB	University	Social interactions: (a) video modeling (b) verbal cues	Social engagement: (a) eye contact (b) facial expressions (c) turn taking (d) sharing emotions
McMahon et al. (2016)	MPDB	University (IPSE)	Augmented reality: (a) app training (Aurasma) (b) vocabulary intervention	Correct responses: • number on each 20-item vocabulary assessment
Reed et al. (2016)	MBDP	Community College (IPSE)	Split-page note-taking: (a) shortened content (b) bullet-point recording (c) subtopics vs. details	Note quality: • participant content vs. peer or interventionist content
Siko (2018)[a]	MBDP	Community College	• Self-monitoring prompting • Wearable technology • Software (feed prompts)	• On-task behavior • Academic engagement
Wright et al. (2022)	MPDP; ABAB	University (IPSE)	Systematic implementation of reminders (wearable technology) embedded in a reversal design	Appointment completion: (a) appointment name (location) (b) appointment time and day (c) appointment tasks

Note. MBDB = multiple-baseline design across behaviors, MBDP = multiple-baseline design across participants, MPDB = multiple-probe design across behaviors, MPDP = multiple-probe design across participants, ABAB = reversal design, ABCD = multiple-component design, IPSE = inclusive postsecondary education

Table 2. Participant Demographics

Author (Year)	Age	Gender Identity	Ethnic Identity	Disability
Ashbaugh et al. (2017)	18-25	male (1); female (2)	European (1); European-American (2)	ASD
Detar (2020)	18-22	male (2); female (1)	NR	ASD
Engel (2009)[a]	18-20	male (3); female (1)	NR	ASD (3); ASD/ADHD (1)
Engstrom (2019)[a]	19-20	male (2); female (1)	White (2); Asian-American (1)	ASD
Gunn et al. (2017)	22	female (1)	NR	ASD
Jackson et al. (2018)	18-20	male (2); female (1)	Caucasian (3)	ASD
Koegel et al. (2013)	21-23	male (3)	Caucasian (2); Hispanic (1)	ASD
Koegel et al. (2016a)	20-37	male (3); female (2)	Caucasian (4); Asian-American (1)	ASD
Koegel et al. (2016b)	19-26	male (3)	Caucasian (2); Middle Eastern (1)	ASD
Mann & Karsten (2020)	NR	male (2)	NR	ASD (AS)
Mason et al. (2012)	19-26	male (2)	Caucasian (1); Hispanic (1)	ASD (AS)
McMahon et al. (2016)	19-25	male (1); female (3)	NR	ID; ASD
Reed et al. (2016)	20-23	male (2); female (1)	White (3)	ASD
Siko (2018)[a]	19-22	male (3)	NR	ASD
Wright et al. (2022)	19-21	male (3)	NR	ASD; ID

Table 3. Findings and Social Validity

Author (Year)	Social Validity	Findings
Ashbaugh et al. (2017)	Self-report satisfaction questionnaire completed pre- and post-intervention to examine satisfaction across four domains: (a) overall college experience, (b) overall social experience, (c) number of social activities attended, and (d) availability of campus social events. Participants indicated meaningful gains beyond behavioral data.	Participants increased their number of community-based social events, extracurricular activities, and peer interactions
Detar (2020)	Post-intervention interviews were held for each participant, with results indicating further collateral gains and clinical benefits for future research. Participants cited increased peer relationships, increased social engagement, and increased confidence in both an academic and nonacademic context.	Participant improvements noted in (a) total questions asked, (b) conversational pauses, (c) nonperseverative discussion, (d) conversational reciprocity, and (e) perceived confidence
Engel (2009)[a]	Surveys held after the intervention (i.e., end of semester) to determine participant responsiveness to the therapy dog and perceptions of the semester.	Animal-assisted consultations did not indicate significant changes to repetitive behaviors, appropriate verbalizations, attention on the consultant, or attention on the animal
Engstrom (2019)[a]	Measurement at pre- and post-intervention to determine effects on quality of life. No observed increases or decreases among participants at any phase.	Participants saw improvement in the percentage of completed daily living skills
Gunn et al. (2017)	No social validity measures or observations reported.	Participant demonstrated variability in scanning, engagement, and verbal interactions in the classroom but the quality of the interactions increased over time
Jackson et al. (2018)	Pre- and post-intervention surveys to assess participants' use of writing strategies and the benefits of those strategies.	Participants showed improved writing performance across the intervention, including generalization probes throughout each phase

Koegel et al. (2013)	Social validity measurements included evaluating quality of (a) social interaction, (b) academic achievement, and (c) employment activity. Social validity of treatment assessed satisfaction of socialization across four domains: (a) satisfaction with college experience, (b) satisfaction with peer interactions, (c) confidence in peer conversations, and (d) satisfaction with number of friends.	Participants increased their number of weekly social events and reported higher quality of life and satisfaction with peer interactions and their college experience; improvements were also seen in untargeted areas (e.g., non-structured social interactions, grade point averages, employment)
Koegel et al. (2016a)	Social validity of the intervention was assessed using a social desirability scale with coders naïve to the study.	Participants showed a decrease in negative statements with simultaneous increases in positive and neutral statements. Additionally, some participants improved their affect and interest, including an increase in social desirability. Secondary effects included improvements in depression, anxiety, and hopelessness.
Koegel et al. (2016b)	Pre- and post-intervention self-report questionnaires to evaluate participants' confidence in communication and socialization across (a) confidence in peer conversations, (b) confidence in ability to express empathy with peers, (c) confidence in ability to ask questions to peers, and (d) importance of expressing empathy.	Participants improved in verbal expression of empathetic statements and questions in peer conversations during generalization and maintenance phases. Supplemental assessments suggested improvements in general levels of empathy and confidence in communication skills.
Mann & Karsten(2020)	Neurotypical peers naïve to the study rated the awkwardness of videotaped conversations across patterns of disinterested behaviors. Self-report questionnaires were completed by participants in relation to their experiences in the study.	Participants saw an improvement in conversation responses related to interest (i.e., confirm availability, invite contribution, change topic).
Mason et al. (2012)	No social validity measures or observations reported.	One participant saw improvement across all measured domains (i.e., eye contact, facial expression, turn taking, shared emotion), and one participant improved in the eye contact and turn taking domains.

McMahon et al. (2016)	Post-intervention surveys were conducted to assess student perceptions and opinions on the acceptability of the augmented reality intervention, including a Likert scale and open-ended questions.	Participants acquired definitions and labeling knowledge for new science vocabulary.
Reed et al. (2016)	No social validity measures or observations reported.	Participants showed notable gains in note-taking ability and quality of notes, which continued during maintenance probes.
Siko (2018)[a]	Post-intervention interviews were completed by each participant, with questions centering on four primary themes: (1) general perceptions and experiences with the wearable technology, (2) perceived usefulness for monitoring on-task behavior and participating during class, (3) perceived efficacy for monitoring out-of-class academic behaviors, and (4) suggestions related to general supports for students with autism spectrum disorder.	Participant data indicated an improvement in on-task behavior and classroom participation during class lectures but did not show noticeable changes in other academic skills (i.e., organization, time management, task initiation).
Wright et al. (2022)	Participants completed a survey related to perceptions and acceptability of using wearable technology as a self-operated prompting system, including a Likert scale and open-ended questions.	Participants showed a significant increase in independent task initiation and task completion

Table 4. NTACT Quality Indicators for Single-Case Experimental Designs

Study	Design	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Final[a]
Ashbaugh et al. (2017)	MBD	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	**1**
Detar (2020)	MBD	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	**1**
Engel (2009)[b]	MBD	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0	**0**
Engstrom (2019)[b]	MBD	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	**1**
Gunn et al. (2017)	ABCD	1	1	0	1	1	1	1	1	1	1	0	0	0	0	0	0	1	0	1	1	**0**
Jackson et al. (2018)	MBD	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	**1**
Koegel et al. (2013)	MBD	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	**0**
Koegel et al. (2016a)	MBD	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	**0**
Koegel et al. (2016b)	MBD	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	**1**
Mann & Karsten (2020)	MBD	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	**0**
Mason et al. (2012)	MBD	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	**1**
McMahon et al. (2016)	MPB	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	**1**
Reed et al. (2016)	MBD	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	**1**
Siko (2018)[b]	MBD	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	**1**
Wright et al. (2022)	MPD/ABAB	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	**1**
Total Meeting	**#**	15	15	14	14	15	15	15	15	15	15	10	14	14	14	14	14	12	14	5	9	**10**
Percentage	**%**	100	100	93	93	100	100	100	100	100	100	67	93	93	93	93	93	80	93	33	60	**67**

Note. NTACT = National Technical Assistance Center on Transition, MBD = Multiple-Baseline Design, MPD = Multiple-Probe Design, ABAB = Reversal Design.

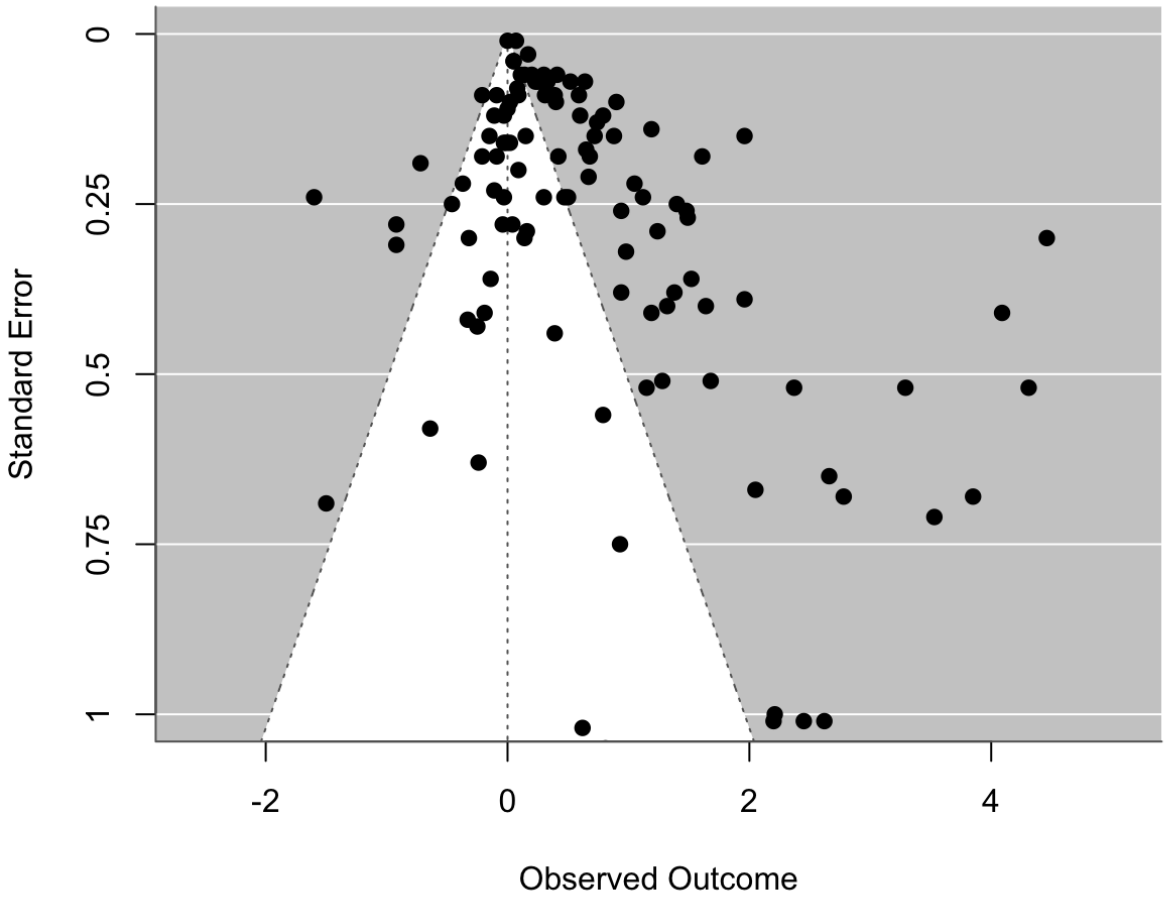
[a] To determine the final ranking, we used the criteria set forth by NTACT. A study rated as high quality must meet all 20 indicators and was coded as 2. A study rated as acceptable quality must meet indicators 1-16 and at least one indicator from 17–20 and was coded as 1. A study failing to meet indicators was coded with a 0.

[b] Denotes dissertations.

Figure 1. Interobserver Agreement (NTACT:C Quality Indicators for SCEDs)

Author	Participants		Setting	Dependent Variable					Independent Variable/Intervention			Procedures		Design/Graphs/Results			Social Validity				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	IOA
Ashbaugh_2017b	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Mann_2020	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.85
McMahon_2016	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0.95
Reed_2016	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0.85
Wright_2022	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0.9
IOA	0.8	0.8	1	0.8	1	1	1	1	1	1	1	1	1	1	1	1	1	0.8	0.6	0.4	

Figure 2. Funnel Plot (Effects Level)



Appendix B: Single-Case Experimental Design (Tables and Figures)

Figure 1. m-Path Daily Prompts (Screenshot)

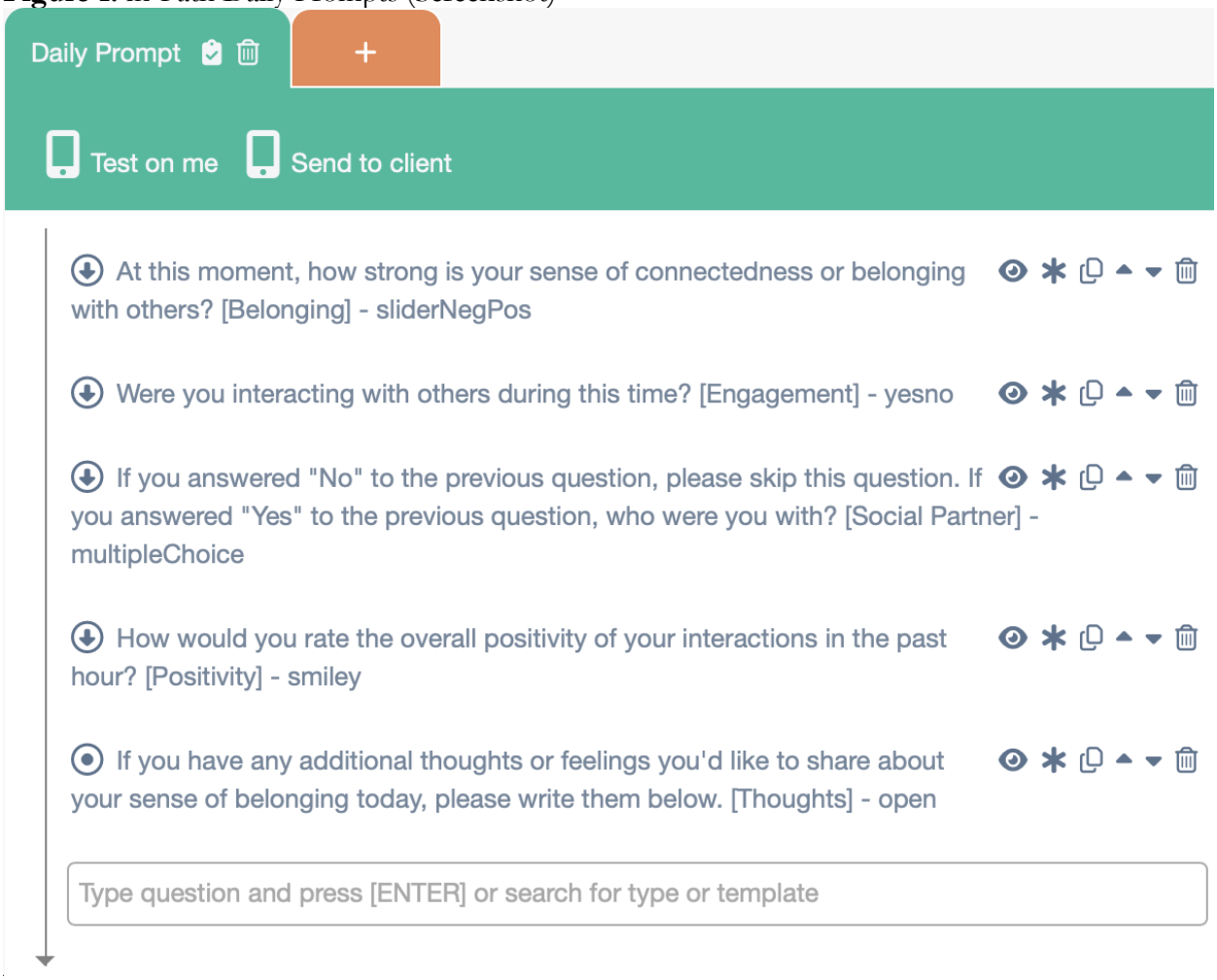


Figure 2. Fidelity Checklist (Chase [Top]; Alyx [Bottom])

	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	Session 7
Study Week	4	5	6	7	8	10	11
Date	2/14/25	2/21/25	2/28/25	3/5/25	3/12/25	3/26/25	4/9/25
- Start Time	9:08	9:12	8:59	12:20	12:44	12:42	1:02
- End Time	9:45	9:35	9:30	12:44	1:12	1:14	1:35
Attendance							
- Participant	X	X	X	X	X	X	X
- Researcher	X	X	X	X	X	X	X
- Observer	X	X	X	X	X	X	X
Intervention							
- Homework Review	N/A	X	X	X	X	X	X
- Participant Practiced Homework	N/A	X	X	X	X	x	X
- Didactic Lesson	X	X	X	X	X	X	X
- Role Play Demonstration	X	X	X	X	X	X	X
- Behavioral Rehearsal	X	X	X	x	X	X	X

	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	Session 7
Study Week	6	7	8	10	11	12	13
Date	2/27/25	3/6/25	3/13/25	3/27/25	4/10/25	4/17/25	4/24/25
- Start Time	1:23	1:18	1:22	1:15	1:24	1:32	1:23
- End Time	2:06	2:02	2:00	2:07	2:00	2:12	2:03
Attendance							
- Participant	X	X	X	X	X	X	X
- Researcher	X	X	X	X	X	X	X
- Observer	X	X	X	X	X	X	X
Intervention							
- Homework Review	N/A	X	X	X	X	X	X
- Participant Practiced Homework	N/A	x	X	X	x	X	X
- Didactic Lesson	X	X	X	X	X	X	X
- Role Play Demonstration	X	X	X	X	X	X	X
- Behavioral Rehearsal	X	X	X	X	X	X	X

Note: *In these sessions, the capital "X" indicates attendance of the individual and completion of the task, whereas a lowercase "x" represents an attempt but not fully completed. "N/A" or "not applicable" is used for the first session since Homework had not been introduced.*

Figure 3. Chase's Belonging (Daily Set)

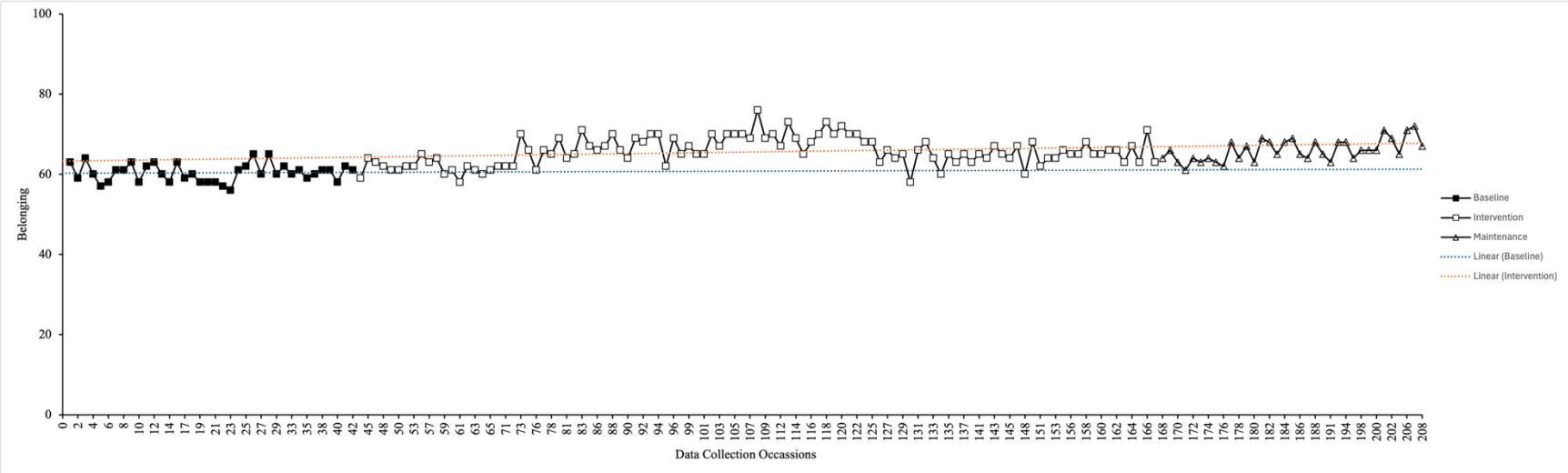


Figure 4. Chase's Positivity (Daily Set)

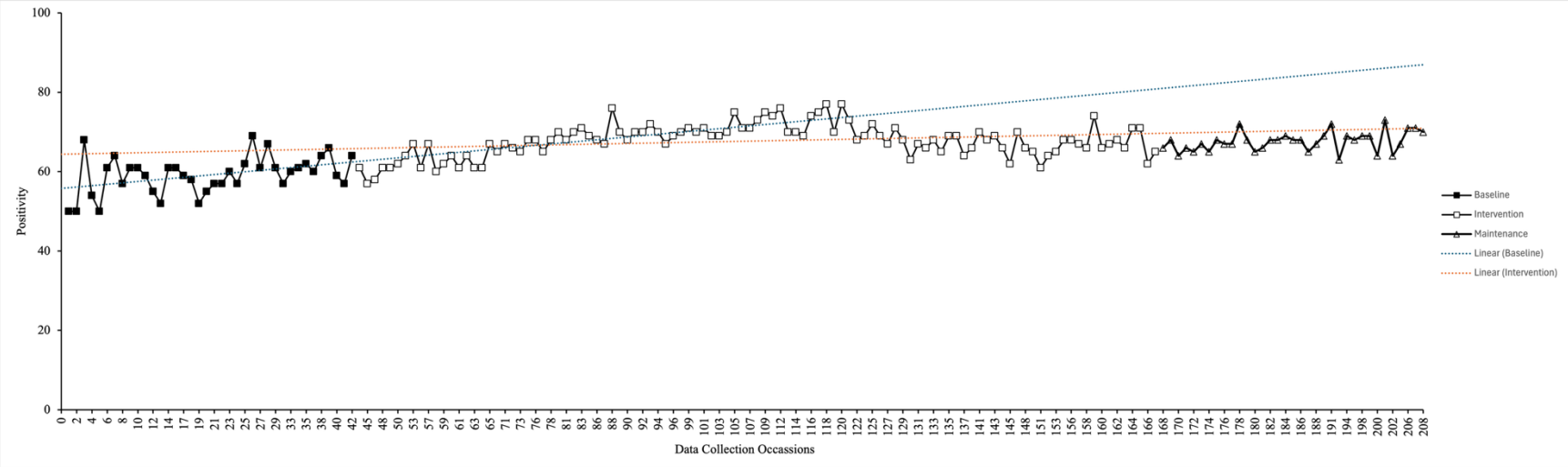


Figure 5. Alyx's Belonging (Daily Set)

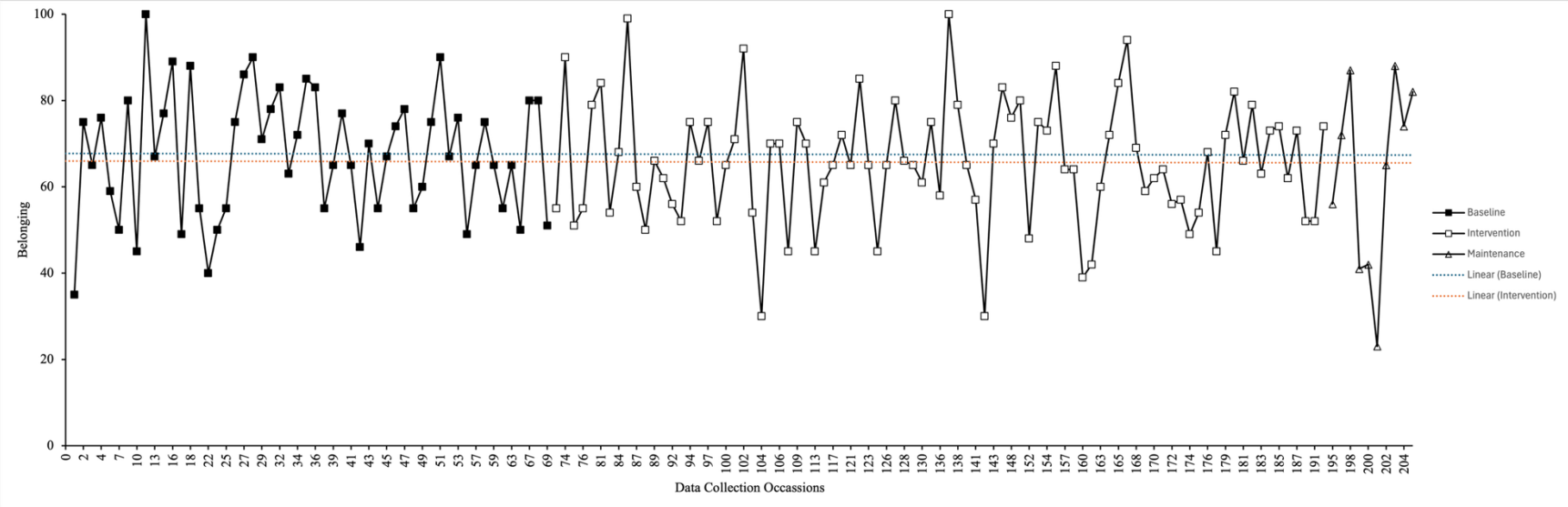


Figure 6. Alyx's Positivity (Daily Set)

