Aided by Extant Data: The Effect of Peer Mentoring on Achievement for College Students with Disabilities

Allison Lombardi[1]
Graham G. Rifenbark[1]
Jessica Monahan[2]
Emily Tarconish[1]
Christopher Rhoads[1]

Abstract

While peer mentor programs for students with disabilities in higher education are increasing in prevalence, the research examining the effectiveness of these programs remains limited. In this study, the effect of one college peer mentoring program on academic performance at a small, private four-year university was examined. The sample was drawn from all students registered with the Office of Disability Services (n = 287), some of whom participated in a peer mentoring intervention as well as a comparison group comprised of non-participants. In light of the observational nature of the data, propensity score weighting and matching were used to adjust for possible confounding variables and to explore robustness to different methodological approaches. Logistic and linear regression methods were used to examine the effect of peer mentoring on academic probation status and grade point average (GPA), respectively, while incorporating propensity score-based adjustments. There were no significant differences between the intervention and comparison groups for either outcome; however, intervention group students had a statistically significantly higher number of accommodations available to them. The study illustrates that extant data, when used in conjunction with appropriate statistical adjustments, is a viable alternative to randomized studies. Implications for higher education researchers and practitioners are discussed, including suggestions to collect various types of non-academic data (e.g., satisfaction, well-being, self-determination surveys) as well as examine structural factors of the program (e.g., mentor-mentee relationships, mentor training) in order to better understand the possible benefits of peer mentor programs and the need for collaborative partnerships between disability services, student affairs, researchers, and institutional research staff.

Keywords: peer mentoring, college students, disability services, propensity score matching, propensity score weighting, quasi-experimental design

As institutions of higher education enroll growing numbers of students with disabilities (SWD), there is an intensified need to develop support programs for these students. Encompassing approximately 11% of undergraduate students, SWD earn fewer credits and are less likely to complete degrees than their peers without disabilities (41% vs. 52%) (Newman et al., 2011). Although they leave postsecondary education for reasons that may be similar to their peers without disabilities (e.g., cost, poor grades, transferring out, health demands and/or family demands), some of these issues may be associated with or exacerbated by disability-related complications (e.g., not using accommodations, increased health or medical issues). As such, college SWD may face multiple barriers that are similar to their peers without disabilities, yet they also have unique disability-related needs that may further compound barriers to successfully completing college degree programs.

While institutions of higher education continue to develop resources to support the unique needs of SWD, research conducted in college settings should inform these practices. Peer mentoring programs, an emerging resource for college students, may be a promising support for SWD. The purpose of this study was twofold: (a) to examine the effectiveness of a peer mentor program for college SWD, and (b) to employ propensity score weighting and matching to extant institutional data. Findings show the promise of using extant data in conjunction with advanced methods for extracting causal estimates from such data as a viable alternative to randomized control group designs in college settings.

[1] *University of Connecticut;* [2] *University of Delaware*

College Peer Mentor Programs

Colleges and universities continue to utilize peer mentor models to offer a naturalistic form of support for students (Budge, 2006). These programs aim to connect a less experienced student with a more advanced individual to promote academic and personal growth. Capitalizing on the experience of upperclassmen, this approach connects novice learners with mentors who are also approachable and relatable (Collier, 2015). Peer mentors can support diverse aspects of a mentee's development, including academic, cognitive, psychological, motivational, or social (Izzo & Shuman, 2013; Jones & Goble, 2012). As both an elder and a peer, a mentor has a unique opportunity to serve as not only a coach but also as a trusted friend, student advocate, and connection to campus resources (Colvin & Ashman, 2010). Forming both personal and reciprocal relationships with their mentees, peer mentors may offer an alternative to fulfill the complex and personal needs of students, especially those undertaking their first year of study (Crisp & Cruz, 2009; Nora & Crisp, 2007; Ward et al., 2014).

Peer Mentor Programs and Outcomes

Despite the potential promise of these supports, peer mentoring programs in higher education have rarely been examined using rigorous quantitative methods. While few in number, several preliminary quantitative studies at large public universities suggest that active participation in peer mentoring programs may have a positive relationship with academic outcomes (Hryciw et al., 2013; Rodger & Tremblay, 2003). In a randomized control study, Rodger and Tremblay (2003) observed that students who were assigned to and actively worked with peer mentors achieved significantly higher grades than a control group who did not work with mentors. In a non-experimental study of a peer mentor program, Hryciw and colleagues (2013) also found an improvement in grades of students who participated in peer mentoring. Other non-experimental studies have produced qualitative reports from peer mentored students who indicated that their learning had been improved by working with a peer mentor (Fox & Stevenson, 2010; Hryciw et al., 2013; Smith, 2007). Together these findings show emerging evidence of the effects of peer mentor programs for college students, yet quantitative study designs that imply causal inference remain sparse in the literature.

Peer Mentor Programs and Subpopulations

Peer mentor programs have promising results for other minority student groups, such as students of color, and the LGBT community (Budge, 2006). While these different student subpopulations do not necessarily include SWD, these findings suggest peer mentoring could be applicable to SWD as they may face similar barriers or stressors upon entering higher education. One phenomenological study found that Latino/a students benefitted from sharing common experiences and challenges with older student mentors with regard to learning to network and building positive relationships, and several students in this same study conveyed that without the mentoring program, they may not have continued at the university (Salas et al., 2014). Oaks et al. (2013) examined a mentoring program that focused on leadership development of African American male college students. Although the program involved additional elements, such as adult mentoring, qualitative interviews revealed that the students felt their interpersonal skills improved through peer mentoring. Participants specified that peer mentoring provided psychosocial support and feedback and served as a valuable learning platform.

Preliminary research assessing mentoring programs for college students who identify as LGBT have also suggested favorable outcomes. Renn (2007) and Renn and Bilodeau (2005) described qualitative studies investigating student identity, leadership, and activist development at three institutions in the Midwest. Both studies found that in most instances, students identified peer mentoring experiences, both formal and informal, to be an integral part of their development particularly with regard to locating LGBT-affiliated groups on campus, which in turn guided them to take on formal leadership positions within these communities and create lasting friendships. Given the positive experiences of other student subpopulations, research on peer mentor programs should continue with an aim to explore the impact on SWD. While SWD are a growing subpopulation in college settings, these students do not achieve comparable academic outcomes to their peers without disabilities (Newman et al., 2011); this discrepancy highlights the need to identify effective practices to support these learners.

Peer Mentor Programs for Students with Disabilities

Peer mentoring programs for SWD are becoming more prevalent on college campuses, although few studies have assessed these programs' efficacy. Brown et al. (2010) conducted a systematic literature review on peer mentor programs for SWD, where the majority were students with learning disabilities and attention deficit hyperactivity disorder (ADHD). Common themes identified included the use of tech-

nology to connect mentors and mentees, self-reported increases in motivation, time management, and attitude, as well as one study that noted students involved in mentoring identified a decrease in academic anxiety. Other studies found that postsecondary SWD who participated in peer mentoring programs self-reported satisfaction and an increase in their social skills and self-efficacy (Ames et al., 2016; Zwart & Kallemeyn, 2001).

Broadly, studies exploring effective practices in higher education and disability, including peer mentoring, lack enough rigor to make causal inference. In a recent systematic literature review examining articles published in the *Journal of Postsecondary Education and Disability* from 1983 to 2012, findings show that only six studies utilized a control or comparison group (Faggella-Luby et al., 2014), a critical feature of a quantitative design that allows for causal inference (Odom et al., 2005). While these findings are based on articles published in a single journal, it should be noted this journal is the only one of its kind that focuses on research supporting college SWD. Additionally, researchers studying the prevalence and depth of disability-related studies across multiple higher education journals found a limited number of rigorous quantitative research designs (Dukes et al., 2017; Madaus et al., 2018). This limitation is a concern mainly because rigorous quantitative research methods will allow for identification of evidence-based practices in higher education settings, and this gap in the literature hinders efforts to improve outcomes for college SWD.

While some research studies have shown promise in the effectiveness of peer mentor programs, it is imperative to assess causal relationships between peer mentor programs and improved academic outcomes in order to establish evidence-based practices (Odom et al., 2005). As of 2010, only ten studies on this topic showed that peer mentoring could qualify as an evidence-based practice (Brown et al., 2010). Unfortunately, even within those studies, sample sizes were small and at least one (Zwart & Kallemeyn, 2001) reported possible disparities between the treatment and control groups. Researchers must employ causal inference designs to assess the effectiveness of these programs while acknowledging the challenges of implementing random assignment in real-word settings. The purpose of this study is to add to the research base on the effectiveness of peer mentor programs for college SWD. Specifically, the study compares academic outcomes for SWD in a college peer mentoring program to the outcomes of their peers with disabilities not involved in the program at one four-year institution. Propensity score weighting and matching methods were employed in order to demonstrate that rigor can be preserved even in a retrospective study using extant data when appropriate methods are utilized. These analytic approaches illustrate that leveraging extant data readily available through a university data warehouse to study the effectiveness of peer mentoring can be a viable alternative to randomized designs.

Method

Participants

Participants were students who attended a small, private university in New England from Fall 2014 through Spring 2016. The entire sample was comprised of 287 SWD; of which 46 (16%) were mentees in the Peer Mentoring program and served as the intervention group, and 241 were comparison students who did not participate in the Peer Mentoring program. Conditioning on the full sample, 39% had overlapping disabilities (i.e., comorbid), the most prevalent overlapping disability category observed was ADHD (40%), followed by LD (39%); with 3.52 mean of accommodations (SD = 2.16). The majority of the SWD were Male (63%) and White (88%). Table 1 contains detailed sample characteristics.

The Peer Mentor Program

The peer mentor program (PMP) was the intervention in the current study. The PMP matched trained undergraduate mentors with students registered with Student Disability Services to provide social support and guidance, and was overseen by Student Disability Services staff, psychology department faculty members, and Ph.D. students in the Behavior Analysis program. Mentors included sophomore, junior, or senior psychology majors who wanted to learn more about peer and social supports. After interviewing with Student Disability Services, mentors attended two 1.5-hour supervision sessions with mentor supervisors and current mentors to complete initial training requirements, including an introduction to Applied Behavioral Analysis, the services offered by Student Disability Services, how mentors can assist students, and the types of issues mentees may experience. Training also included role-playing so that new mentors could practice responding to different situations that may arise when working with mentees.

The PMP primarily served students with autism spectrum disorders, ADHD, or executive functioning impairments. PMP is introduced to students at freshmen orientation or during their intakes with Student Disability Services before they arrive at the university. As such, most mentees started working with peer mentors as freshmen. Both mentors and mentees

committed to participate with the program for at least one year, and were able to continue until graduation.

Mentors typically met with mentees for one hour per week, sometimes for one long session, two half hour meetings, or four 15-minute meetings; meeting times were adjusted to the needs of the mentee. During the first interaction, mentors helped mentees define the measurable goals they wanted to accomplish and created plans to track success. Subsequent sessions focused on practicing goal-related tasks and reviewing mentees' progress.

Mentors were encouraged to establish a plan for each meeting. Typical activities included reviewing and planning future homework and study strategies, helping mentees to schedule their time, or practicing conversation or social interaction strategies. Mentors could suggest different ways to introduce oneself into a conversation, how to assess when it is appropriate to enter a conversation, or ways to get involved on campus, depending on the mentees' areas of difficulty. As many mentees struggled with peer interactions, the act of having a time and space to rehearse social skills was beneficial (Ward et al., 2014). Mentors maintained contact logs for every appointment, which they shared with Student Disability Services during weekly supervision.

Procedures

The gold standard for making causal inferences about the average effectiveness of a program is to utilize an experimental design, whereby individuals either receive the treatment (i.e., are placed into the program), or are placed into the control group (i.e., are not placed into the program) as the result of an entirely random process. Thus, in a randomized study the individuals receiving treatment are just as likely to have been placed in the control group (i.e., not placed in the program). In the event such a randomized design is employed, treatment assignment is strongly ignorable in the sense that, on average, those receiving the program and those not receiving it are balanced with respect to all possible covariates (e.g., race, gender, socioeconomic status). In other words, selection into the treatment group does not depend on an individual's race, gender, socioeconomic status, or any other measured or unmeasured variable. However, random assignment to treatment is not always feasible. In the current study the decision to participate in the PMP program or not was made by individual students rather than a random process. Therefore, the current study is an observational study rather than a random experiment.

Inferences about program effectiveness based on observational data require statistical adjustments to account for pre-existing differences between the treatment and comparison groups. Without adjustments, imbalances on observed covariates (e.g., demographic attributes) can confound inferences about the causal effects of the program of interest. One such example is the PMP in this study, since students in the intervention group self-selected into the program, while students in the comparison group, chose not to participate. All study protocols were reviewed and approved by Institutional Review Boards of the institution at which the intervention took place as well as the researchers' institution.

Measures

All measures were in the form of extant data that universities typically collect and might be housed in a university data warehouse, such as demographics, scores on higher education entrance exams (i.e., SAT); academic data (i.e., number of credits attempted and earned, probation, suspension); and attrition data (i.e., expulsion, graduated, withdrawn).

Covariates. Our propensity score-based procedures aim to achieve balance between the treatment and control groups with respect to the following covariates: gender, minority status, ADHD diagnosis, LD, mental health, and ASD; as well as comorbidity (i.e., multiple diagnoses).

Outcomes. Academic data were available at four separate time points that spanned from the Fall semester of 2014 through the Spring semester of 2016. Of interest was the effect of the PMP on academic standing (i.e., academic probation) and grade point average (GPA). Different patterns of student participation in the PMP created difficulties in determining how to define our treatment indicator variable and which measurement(s) to use to assess outcomes. A total of 8 patterns surfaced. The most common was participation in all 4 semesters, however, this pattern still represented only 30% (n = 14) of the intervention group sample. Our decisions regarding how to define treatment and outcome measurement are described in more detail below.

Academic Probation. Intervention students were coded as on academic probation if they were in poor academic standing at the end of their final semester in the PMP. This decision rule allowed the success of the PMP to be determined based on final status of the mentoring, instead of the student's academic standing while actively participating. In terms of comparison students, we coded them as being on academic probation, if they were ever reported to be in poor academic standing during the semesters in which we observed them. In sum, these coding decisions led to 13% of the intervention students (n = 6) and 17% of the comparison students (n = 40) to be coded as on academic probation.

Grade Point Average (GPA). The observed GPA for an intervention student in their final semester in the PMP was used as the outcome; whereas, comparison student outcomes were taken as their mean GPA across all semesters from which this data was available. The mean GPA for intervention students was 2.524 (SD = 0.988) and 2.894 (SD = 0.722) for comparison students.

Data Analysis

In this study, we address the non-random selection into the PMP by employing a propensity score weighting and a matching approach. Then, regardless of approach (weighting or matching), the effect of PMP on academic probation was determined via logistic regression – which estimates the log odds of being placed on academic probation; whereas, the effect of PMP on grade point average was estimated using an ordinary least squares regression – which provides an estimate of the average difference between PMP participants and non-participants with respect to GPA. All statistical analyses were conducted in R (R Core Team, 2017).

Propensity Score Weighting. In order to execute a PSW approach, (a) initial bias must be investigated, (b) propensity scores must be estimated and used to generate weights, and (c) the effect of the weights on bias must be assessed. Upon satisfactory reduction of bias, the causal estimate (e.g., the effect of PMP on academic outcomes) can be pursued.

Initial Bias. Logistic regression models are estimated in which each covariate in separate models are regressed onto intervention status. These models are used to determine the degree to which the average value of student attributes differ as a function of intervention or control group. When the exponent of the intervention effect regression estimate ($\hat{b}$) is taken, an *odds ratio* (OR) results. OR can be used as an effect size, where OR values of 1.68, 3.47, and 6.71 correspond to small, medium, and large effect sizes. (Chen et al., 2010). If there is perfect balance, OR will equal 1.0.

Propensity Scores and Weights. In order to generate the appropriate weights, a logistic regression model is estimated in which intervention status is regressed onto all study covariates. From this model, propensity scores are estimated representing the probability that an individual with a given covariate vector belongs to the intervention group. Using the estimated propensity score and the observed intervention status (Int.Status) for each individual, weights for individual i can be generated (see below) that corresponds to the estimand of interest: the *average treatment effect on the treated* (ATT).

$$\text{ATTi} = \frac{\text{Int. Status}_i}{1} + \frac{(1\text{-Int. Status}_i) * (\text{PS}_i)}{(1\text{-PS}_i)}$$

When estimating the ATT individuals in the intervention group (Int.Status = 1) are all given a weight of 1. On the other hand, since individuals with larger propensity scores are more likely to be in the intervention group, the comparison group case with higher propensity scores (i.e. those that are more similar to their intervention counterparts) receive larger weights. For instance, an individual with a propensity score of 0.8 would get an ATT weight of 4 whereas an individual with a PS of 0.4 will be given an ATT weight of 0.667.

Re-evaluate Bias. To assess the reduction in bias, we estimated separate weighted logistic regressions in which study covariates were regressed onto intervention status, using the ATT weights. By utilizing these weights, it is possible to determine the degree to which bias has been reduced by allowing some comparison students to be more influential than others. After we achieved satisfactory balance on the study covariates, weighted regression models were estimated for each of the outcome variables. These regression models also included all study covariates.

Matching. Matching is a preprocessing approach that helps facilitate causal claims. By utilizing matching in addition to regression adjustment the causal claims resulting from a statistical model are doubly robust and likely will not result in bias (Ho et al., 2011). Possible approaches to matching are one-to-one and one-to-many matching in which each intervention student is matched with one or more than one comparison student, respectively. In fact, omitting observations can help reduce bias when there are comparison students that are very dissimilar from all students in the intervention group. Due to the nature of matching methods (i.e., the outcome variable never utilized) it is appropriate to investigate many matching methods to determine which method best achieves the dual goals of: (a) ensuring the greatest balance between groups and (b) maximizing the precision of the estimate of the ATT

Assessing Balance. Balance is assessed in the same manner as when propensity scores are used as weights. Using the full sample, initial differences between intervention and control groups with respect to the covariates is investigated. After matching, these differences are re-estimated using the matched sample. The difference in these estimates represents the percent improvement in balance. Improvement in balance was also assessed by generating QQ plots for both the unmatched and matched sample. We elected

to generate a one-to-one matched sample using the nearest neighbor approach available in the MatchIt (Ho et al., 2011) R package.

Results

Results are presented in two parts: (a) propensity score weighting, and (b) propensity score matching. Within each part, the effect on study outcomes academic probation and GPA are described.

Propensity Score Weighting

Due to observed missing values on the minority covariate, 18 observations were removed from the comparison group prior to conducting the analyses, leaving an overall sample size of 271 (comparison: $n = 225$; intervention; $n = 44$).

Initial Bias. A total of seven logistic regressions were estimated to investigate initial bias. With respect to student demographics, we found that intervention status had an effect on gender, specifically, Males were estimated to be 7.84 times more likely in PMP ($\hat{b} = -2.06$, $SE = 0.54$, $p < 0.001$; *OR*: 7.84 = [1/0.13]). Whereas, intervention status did not have an effect on minority status ($\hat{b} = -0.48$, $SE = 0.51$; $OR = 0.62$).

With respect to disability categories, we found that those diagnosed with ADHD were 4.76 times more likely to be in PMP ($\hat{b} = 1.56$, $SE = 0.35$, $p < 0.001$; $OR = 4.76$), while those diagnosed with ASD were 9.77 times more likely to be in PMP ($\hat{b} = 2.28$, $SE = 0.44$, $p < 0.001$; $OR = 9.77$). The other disability diagnoses such as LD ($\hat{b} = 0.37$, $SE = 0.33$; $OR = 1.45$) and Mental Health ($\hat{b} = 0.03$, $SE = 0.37$; $OR = 1.02$) were non-significant. With respect to those with multiple diagnoses (i.e., comorbidity) they were found to be 2.85 times more likely to be in PMP ($\hat{b} = 1.05$, $SE = 0.33$; $OR = 2.85$).

Generating Weights. Intervention status was regressed onto all study covariates in order to utilize the model predicted probabilities (i.e., propensity of a student being enrolled in PMP) to generate an ATT weight for each student. Using the aforementioned formula, the intervention group had a mean of 1 with standard deviation of 0; whereas, the comparison group had a mean of 0.21 and a standard deviation of 0.57 on the ATT weight variable.

Re-Assess Balance. Using the ATT weights generated, we estimated weighted logistic regressions to investigate whether or not bias was reduced on the study covariates. We found that Males were 1.16 times more likely to be enrolled in PMP ($\hat{b} = -0.15$, $SE = 0.71$, $p = 0.836$; $OR = 0.86$), while minority status remained non-significant ($\hat{b} = 0.03$, $SE = 0.69$, $p = 0.96$; $OR = 1.03$). With respect to diagnoses, we found the effect of intervention status on ADHD ($\hat{b} = -0.15$, $SE = 0.47$, $p = 0.75$; $OR = 0.86$) and ASD ($\hat{b} = -0.13$, $SE = 0.44$, $p = 0.77$; $OR = 0.88$) to be non-significant; while the other disability categories remained non-significant. Finally, the effect of intervention status on multiple diagnoses was found to be non-significant ($\hat{b} = -0.15$, $SE = 0.43$, $p = 0.74$; $OR = 0.88$). Upon these nil findings, bias was satisfactorily reduced using the ATT weights generated. Table 2 shows estimates from both the unweighted and weighted logistic regressions for all study covariates.

Academic Probation. For the unweighted model, we estimated a logistic regression in which academic probation was regressed onto intervention status and all study covariates. We found the conditional log-odds of the intervention effect to be significantly different from 1.0 at the 0.1 level ($\hat{b} = -1.073$, $SE = 0.55$, $p = 0.0504$; $OR = 0.34$). Among the covariates, the conditional log-odds for gender ($\hat{b} = -1.07$, $SE = 0.44$; $OR = 0.34$) and ADHD ($\hat{b} = 1.14$, $SE = 0.44$; $OR = 3.11$) were significantly different from one. For the weighted model, using the ATT weights, the logistic regression was re-estimated to evaluate an unbiased estimand of the effect PMP has on academic probation. Upon fitting this model, we found the effect of PMP to remain non-significant ($\hat{b} = -0.86$, $SE = 0.58$; $OR = 0.42$), as were all study covariates.

Grade Point Average. For the unweighted model, after estimating a linear regression in which GPA was regressed onto intervention status and all study covariates, we found that the effect of PMP was non-significant ($\hat{b} = -0.11$, $SE = 0.13$, $p = 0.394$), therefore, controlling for demographics and diagnoses, those in the PMP do not perform better than their counterparts. The model implied mean GPA for those not enrolled in the PMP (e.g., Caucasian Males) was estimated to be 3.03 on average ($\hat{b} = 3.03$, $SE = 0.09$, $p < 0.001$). Regarding the performance of study covariates, Females were estimated to have a higher conditional GPA on average than Males ($\hat{b} = 0.38$, $SE = 0.09$, $p < 0.001$); whereas, those diagnosed with ADHD ($\hat{b} = -0.56$, $SE = 0.11$, $p < 0.001$), *LD* ($\hat{b} = -0.26$, $SE = 0.11$, $p < 0.05$), and Mental Health ($\hat{b} = -0.34$, $SE = 0.13$, $p < 0.01$) were estimated to have a lower GPA on average. All other covariates were non-significant.

For the weighted model, by employing the ATT weights, the population estimate for the PMP remained non-significant ($\hat{b} = 0.01$, $SE = 0.10$, $p = 0.92$). After weighting, the expected GPA for those in the comparison group in the reference categories (e.g., Caucasian, Male) was estimated to be 3.21 ($\hat{b} = 3.21$, $SE = 0.14$, $p < 0.001$). With respect to study covariates, the effect of ADHD ($\hat{b} = -0.81$, $SE = 0.14$,

$p < 0.001$) and LD ($\hat{b} = -0.38$, $SE = 0.14$, $p < 0.001$) remained significant; whereas, the effect of ASD and gender were non-significant.

Matching

Due to restricting the sample to Caucasian Males, a total of 144 observations (comparison $n = 109$; intervention $n = 35$) were available for the purposes of creating a matched sample. When using the nearest neighbor approach to identifying a matched sample, we settled on the greedy mechanism, whereby each intervention student is matched with the single best comparison (i.e., the shortest distance – on the logit scale) student. This effort resulted in analytic sample of 70 students equally split ($n = 35$, each) between comparison and intervention groups.

Percent Balance. In order to determine the performance of the matching procedure, the mean difference on the study covariates between groups are estimated separately using both the unmatched and matched sample. If the matching procedure performed well, it is expected these estimates will become smaller on an absolute scale. We found that the largest percent reduction occurred for the Other disability category (91.07%), whereas, the smallest percent reduction occurred for the ASD disability category (30.02%). In sum, the mean percent reduction was 75.76 ($SD = 25.81$). Table 3 shows the percent reduction estimates.

Academic Probation. After the reduction in bias in the study covariates, the effect of PMP remained non-significant ($\hat{b} = -0.65$, $SE = 0.61$, $p = 0.289$) leading to a conditional odds of 1.91. All other study covariates were non-significant.

Grade Point Average. Upon regressing GPA onto intervention status and all study covariates, we found that those in the comparison group with a Mental Health disorder were expected to have a GPA of 2.56 ($\hat{b} = 2.56$, $SE = 0.38$, $p < 0.001$) on average. All other terms in the model, including intervention status, were not significantly different from zero. Therefore, we find that students in PMP do not have, on average, higher GPAs than their counterparts ($\hat{b} = -0.15$, $SE = 0.22$, $p = 0.511$).

Discussion

Currently, what is known about college students with disabilities and peer mentor programs is mostly descriptive and qualitative in nature. While there is some emerging evidence to show the promise of effectiveness (Brown et al., 2010), evidence of effectiveness based on causal inference of such programs is largely absent from the higher education literature. The current study sought to address this gap in the literature by applying propensity score weighting and matching methods to extant data in order to compare students who participated in a peer mentor program with their peers with disabilities who elected not to participate. Importantly, this approach demonstrates the possibility of designing a rigorous enough study so that causal inference can be determined but without the need to randomly assign participants into treatment and control groups, a feat that may seem close to impossible in applied research settings.

The results of the current study show no significant effects of peer mentoring on academic achievement and standing using both weighting and matching methods. This finding is contrary to a previous experimental study on the effects of peer mentoring on similar academic outcomes (Rodger & Tremblay, 2003). Despite this difference, there are generally very few experimental studies that examine the effects of peer mentoring on academic achievement. As such, it is not surprising to find mixed results, which further demonstrates the need to prioritize peer mentor studies in higher education settings. Interestingly, we found that intervention group students had a statistically significantly higher number of accommodations available to them. Further investigation of this finding was outside of the scope of the current study; however, this finding suggests that peer mentor program participants may be better connected with the campus disability services office, and simply may have more knowledge of and a tendency to use available supports.

While the results of this study show little effects on academic outcomes, further investigation of important non-academic outcomes are needed to better understand and clarify the utility and potential benefits of such programs. For example, previous study findings showed that peer mentored students with both high and low levels of anxiety achieved comparable grades; whereas, in the control group, students with high anxiety performed worse than those with low anxiety (Roger & Tremblay, 2003). Further, in other non-experimental studies, students not receiving peer mentoring also experienced decreases in self-esteem, and perceived social support, (Collings et al., 2014; Hryciw et al, 2013.) Peer mentored students in this program also indicated experiencing higher levels of integration into the university, whereas, students who did not work with a mentor were four times as likely to indicate wanting to leave the university (Collings et al., 2014). While mentoring programs seem to produce potential benefits for a general student population, their effects may be even more pronounced for specific subpopulations of students.

Limitations

A major limitation to consider in the interpretation of the current study findings is the sample size. The sample was small and limiting, taken from a private four-year college, and based on the students who decided to register with the campus office of disability services and voluntarily participated in the PMP. Among the sample, the participants in the PMP were primarily white males. As such, findings should not be generalized but should be considered in the design of future studies involving campus peer mentor programs. Specifically, even though this study was limited to a single institution, the concept of applying propensity score weighting and matching was demonstrated and potentially could be applied to multi-institutional studies on a much larger scale.

Another important limitation to consider is the approach we took to determining the outcomes for intervention and comparison students. Due to the nature of the data, we had to make decisions about how to treat probation status and GPA that was slightly different between groups (see p. 148). In future studies, these are critical decisions that should be carefully considered with regard to the generalizability of the findings.

Implications for Research and Practice

The results of this study highlight the fact that success of peer mentor programs may not be measured by short term academic variables (GPA and academic probation status) alone. In light of these results, along with previous research in college peer mentor programs, higher education professionals involved in these programs should make intentional decisions about data collection at the beginning, during, and end of a student's experience in the peer mentor program. Also, it is important to consider collecting non-academic data that is potentially more proximal in nature so that academic and nonacademic benefits of peer mentor programs can be clarified. Finally, our post hoc findings on accommodation use were informative and suggest that intervention students were more informed of available supports than those who did not participate in the PMP. Further exploration as to how peer mentor and other types of formal and informal supports affects accommodation use is warranted. It may also be beneficial to consider disability type as well as mentor-mentee relationship dynamics and mentor training in future studies as contributing factors.

With regard to non-academic data collection, self-report surveys may be a viable strategy. For example, programs may consider utilizing self-determination assessments to explore growth in areas such as advocacy, goal setting, or self-efficacy. While research in self-determination in higher education settings is limited, high school students who are self-determined are more likely to experience favorable post-school outcomes (Shogren et al., 2015; Wehmeyer & Palmer, 2003). Further, the Association for Higher Education and Disability (AHEAD) recommends encouraging the development of self-determination skills in postsecondary students with disabilities. The Self-Determination Inventory: Student Report Version (SDI-RS) measures self-determination skills for students ages 18 through 22, with and without disabilities (Shogren et al., 2017) and could be an important tool to measure critical nonacademic skills of students in a peer mentor program.

Ultimately, college peer mentoring programs could be a viable support for SWD, but we know very little about the effectiveness of such programs. This study suggests a promising method of program evaluation where extant data obtained from a university data warehouse is utilized to better understand program effects on short-term academic outcomes. While the current study showed no significant effects of the mentor program on academic outcomes, the results inform the higher education literature on next steps to take as far prioritizing research on peer mentor programs, including data collection of proximal non-academic skills, careful attention to mentor-mentee pairing and mentor training, and coordinated efforts from multiple campus personnel across units in Student Affairs, Disability Services, Institutional Research, and faculty who aim to establish research lines in higher education and disability.

References

Ames, M. E., McMorris, C. A., Alli, L. N., & Bebko, J. M. (2016). Overview and evaluation of a mentorship program for university students with ASD. *Focus on Autism and Other Developmental Disabilities, 31*, 27-36.

Brown, S. E., Takahaski, K. & Roberts, K. D. (2010). Mentoring individuals with disabilities in postsecondary education: A review of the literature. *Journal of Postsecondary Education and Disability, 23*, 98-111.

Budge, S. (2006). Peer mentoring in postsecondary education: Implications for research and practice. *Journal of College Reading and Learning, 37*, 71-85.

Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics—Simulation and Computation, 39*, 860-864.

Collier, P. (2017). Peer mentoring: A tool for serving the diverse needs of 21st century college students. *Metropolitan Universities, 28*(3), 3–8.

Collings, R., Swanson, V., & Watkins, R. (2014). The impact of peer mentoring on levels of student wellbeing, integration and retention: A controlled comparative evaluation of residential students in UK higher education. *Studies in Higher Education, 68*, 927-942.

Colvin, J. W., & Ashman, M. (2010). Roles, risks, and benefits of peer mentoring relationships in higher education. *Mentoring & Tutoring: Partnership in Learning, 18*, 121–-134.

Crisp, G., & Cruz, I. (2009). Mentoring college students: A critical review of the literature between 1990 and 2007. *Research in Higher Education, 50*, 525-545.

Dukes, L. L., Madaus, J. W., Faggella-Luby, M., Lombardi, A., & Gelbar, N. (2017). PASSing college: A taxonomy for students with disabilities in postsecondary education. *Journal of Postsecondary Education and Disability, 30*, 111-122.

Faggella-Luby, M., Lombardi, A., Lalor, A., & Dukes, L. L. (2014). Methodological trends in disability and higher education research: Historical analysis of the Journal of Postsecondary Education and Disability. *Journal of Postsecondary Education and Disability, 27*, 357-368.

Fox, A., Stevenson, L., Connelly, P., Duff, A., & Dunlop, A. (2010). Peer-mentoring undergraduate accounting students: The influence on approaches to learning and academic performance. *Active Learning in Higher Education, 11*, 145-156.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software, 42*, 1-28.

Hryciw, D. H., Tangalakis, K., Supple, B., & Best, G. (2013). Evaluation of a peer mentoring program for a mature cohort of first-year undergraduate paramedic students. *Advances in Physiology Education, 37*, 80-84.

Izzo, M. V., & Shuman, A. (2013). Impact of inclusive college programs serving students with intellectual disabilities on disability studies interns and typically enrolled students. *Journal of Postsecondary Education and Disability, 26*, 321-335.

Jones, M. M., & Goble, Z. (2012). Creating effective mentoring partnerships for students with intellectual disabilities on campus. *Journal of Policy and Practice In Intellectual Disabilities, 9*, 270-278.

Madaus, J. W., Gelbar, N. W., Dukes, L. L., Lalor, A. R., Lombardi, A., Kowitt, J., & Faggella-Luby, M. (2018). Literature on postsecondary disability services: A call for research guidelines. *Journal of Diversity in Higher Education, 11*, 133-145.

Newman, L., Wagner, M., Knokey, A. M., Marder, C., Nagle, K., Shaver, D.,… Schwarting, M. (2011). *The post-high school outcomes of young adults with disabilities up to 8 years after high school. National Longitudinal Transition Study-2 (NLTS2) (NCSER 2011-3005)*. SRI International.

Nora, A., & Crisp, G. (2007). Mentoring students: Conceptualizing and validating the multi- dimensions of a support system. *Journal of College Student Retention, 9*, 337-356.

Oaks, D., Duckett, K., Suddeth, T., & Kennedy-Phillips, L. (2013). Leadership development and the African American male college student experience. *Journal of College and Character, 14*, 331-340.

Odom, S. L., Brantlinger, E., Gersten, R., Horner, R. H., Thompson, B., & Harris, K. R. (2005). Research in special education: Scientific methods and evidence-based practices. *Exceptional Children, 71*, 137-148.

R Core Team. (2017). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Renn, K. A. (2007). LGBT student leaders and queer activists: Identities of lesbian, gay, bisexual, transgender, and queer identified college student leaders and activists. *Journal of College Student Development, 48*, 311-¬330.

Renn, K., & Bilodeau, B. (2005). Queer student leaders: An exploratory case study of identity development and LGBT student involvement at a Midwestern research university. *Journal of Gay & Lesbian Issues in Education, 2*(4), 49-71.

Rodger, S., & Tremblay, P. F. (2003). The effects of a peer mentoring program on academic success among first year university students. *Canadian Journal of Higher Education, 33*(3), 1-17.

Salas, R., Aragon, A., Alandejani, J., & Timpson, W. M. (2014). Mentoring experiences and Latina/o university student persistence. *Journal of Hispanic Higher Education, 13*, 231-244.

Shogren, K. A., Wehmeyer, M. L., Little, T. D., Forber-Pratt, A. J., Palmer, S. B., & Seo, H. (2017). Preliminary validity and reliability of scores on the Self-Determination Inventory: Student report version. *Career Development and Transition for Exceptional Individuals, 40*, 92-102.

Shogren, K. A., Wehmeyer, M. L., Palmer, S. B., Rifenbark, G. G., & Little, T. D. (2015). Relationships between self-determination and post school outcomes for youth with disabilities. *The Journal of Special Education*, *53*, 30–41.

Smith, D. L. (2007). The relationship of type of disability and employment status in the United States from the behavioral risk factors surveillance system. *Journal of Rehabilitation*, *73*(2), 32-40.

Ward, E., Thomas, E., & Disch, W. (2014). Mentor service themes emergent in a holistic, undergraduate peer-mentoring experience. *Journal of College Student Development*, *55*, 563-579.

Wehmeyer, M., & Palmer, S. B. (2003). Adult outcomes for students with cognitive disabilities three-years after high school: The impact of self-determination. *Education and Training in Developmental Disabilities*, *38*, 131-144.

Zwart, L. M., & Kallemeyn, L. M. (2001). Peer-based coaching for college students with ADHD and learning disabilities. *Journal of Postsecondary Education and Disability, 15*, 1-15.

About the Authors

Allison Lombardi is currently an Associate Professor in the Department of Educational Psychology and teaches undergraduate and graduate courses in the Special Education program, as well as a Research Scientist with the Center for Behavioral Education and Research, and a Research Associate with the Center on Postsecondary Education and Disability at the University of Connecticut. Her research interests include college and career readiness for students with disabilities and promoting inclusive instruction among university faculty. She received her B.A. degree in English Literature and her M.A. degree in Education from the University of California, Berkeley, and Ph.D. in Special Education from the University of Oregon. She can be reached by email at: allison.lombardi@uconn.edu.

Graham G. Rifenbark received his B.A. degree in Psychology from the University of Kansas (KU) and his Ph.D. in Educational Psychology with a concentration in research methods, measurement, and evaluation from the University of Connecticut. He is currently a Postdoctoral Research Associate in the Department of Educational Psychology at the University of Connecticut. In this position, he serves a methodologist on grants that have been awarded and that are pending. His research interests include latent variable and mixed effects modeling, data-model fit, and randomized/quasi-experimental research designs in the context of secondary special education and transition. He can be reached by email at: graham.rifenbark@uconn.edu.

Jessica Monahan received her M.A. and Ph.D. in Educational Psychology from the University of Connecticut. Her experience includes working as a high school special education teacher, a program coordinator for an inclusive postsecondary education program for individuals with intellectual disabilities, and an instructor at Southern Connecticut State University. She is currently a Research Manager at the Center for Disability Studies at the University of Delaware. Her research interests include college and career readiness for students with disabilities and supports for students with disabilities in postsecondary education. She can be reached by email at: jlmon@udel.edu.

Emily Tarconish received B.A. degrees in English and Women's Studies and an M.S. in Rehabilitation Counseling from The Pennsylvania State University. Her experience includes working as a vocational rehabilitation counselor with The Pennsylvania Office of Vocational Rehabilitation and as the Director of Student Accessibility Services at Clark University. She is currently a PhD student in the Department of Educational Psychology at the University of Connecticut. Her research interests includes the experiences of and services for postsecondary students with disabilities, particularly those with traumatic brain injuries, and inclusive teaching practices for postsecondary instructors. She can be reached by email at: emily.tarconish@uconn.edu.

Chris Rhoads received his B.A. degree in Philosophy from Haverford College and Ph.D. in Statistics from Northwestern University. He is currently a professor in the Department of Educational Psychology at the University of Connecticut. His research is in the area of methods for causal inference from observational studies in educational settings, especially settings with a multi-level structure. He can be reached by email at: Christopher.rhoads@uconn.edu.

Table 1

Sample Characteristics

	Comparison	Intervention	Overall
n	241	46	287
Gender			
Male	58%	91%	63%
Female	42%	9%	37%
Race			
Caucasian	87%	91%	88%
African American	5%	3%	5%
Asian	3%	3%	3%
Other	1%	3%	1%
Multiracial	4%	-	3%
Ethnicity			
Non-Hispanic	9%	92%	91%
Hispanic	10%	8%	9%
Disability			
ADHD	34%	72%	40%
LD	38%	48%	39%
Mental Health	28%	26%	28%
ASD	5%	33%	9%
Other	37%	4%	32%
*Comorbid	34%	61%	39%
Number of Accommodations Mean (SD)	3.39 (2.12)	4.24 (2.25)	3.52 (2.16)

*Indicates multiple disability categories observed.

Table 2

Assessment of Covariate Balance - PSW

Response Variable	Unweighted OR	Weighted OR
Male	7.84*	1.16
Minority	0.618	1.03
ADHD	4.763*	0.86
LD	1.451	0.79
Mental Health	1.029	1.24
ASD	9.773*	0.88
Comorbidity	2.846*	0.88

* Indicates the log-odds were greater than 1.

Table 3

Percent Reduction in Group Mean Difference

	Group Mean Difference
Distance	63.701
ADHD	85.160
ASD	30.016
Other	91.073
Comorbid	90.082
No. Accommodations	82.465