

Clemson University

TigerPrints

All Dissertations

Dissertations

August 2021

The Effectiveness of a Diversity and Inclusion Intervention in Higher Education in the Context Of COVID-19

Katherine Brown D'Souza

Clemson University, knbrown3093@gmail.com

Follow this and additional works at: https://tigerprints.clemson.edu/all_dissertations

Recommended Citation

D'Souza, Katherine Brown, "The Effectiveness of a Diversity and Inclusion Intervention in Higher Education in the Context Of COVID-19" (2021). *All Dissertations*. 2874.

https://tigerprints.clemson.edu/all_dissertations/2874

This Dissertation is brought to you for free and open access by the Dissertations at TigerPrints. It has been accepted for inclusion in All Dissertations by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.

THE EFFECTIVENESS OF A DIVERSITY AND INCLUSION INTERVENTION IN
HIGHER EDUCATION IN THE CONTEXT OF COVID-19

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Industrial-Organizational Psychology

by
Katherine Brown D'Souza
August 2021

Accepted by:
Dr. Patrick Rosopa, Committee Chair
Dr. Mary Anne Taylor
Dr. Jennifer Bailey Bisson
Dr. Sarah Winslow

ABSTRACT

Previous research has shown that diversity in the workplace can have varying effects, depending on how well organizations manage their workforce diversity (Mor Barak et al., 2016). When organizations properly manage their workforce, they can reap the benefits of diversity while avoiding the pitfalls. Diversity training interventions are one way many organizations choose to manage their diverse workforces effectively while creating organizational climates of inclusion (Kalinowski et al., 2013). The present study explored the impact of four different diversity training interventions designed to promote greater awareness of gender and racial bias in academia and to promote bias reduction efforts. Compared to a control group selected using propensity score analysis, trainees from across all programs were more concerned about gender discrimination, perceived a more positive organizational inclusion climate, had more positive attitudes toward workplace diversity, and exhibited lower levels of sexism. The present study also examined the impact of COVID-19 on faculty willingness and ability to participate in voluntary learning opportunities. Faculty who experienced drastic decreases in the time they could dedicate to research during COVID-19 reported being less able and willing to participate in voluntary learning opportunities. Additionally, faculty with significant childcare responsibilities were less able (but not less willing) to participate in voluntary learning opportunities than faculty with no childcare responsibilities. The present study contributed to diversity training research and practice by further examining attitudinal training outcomes, using a novel propensity scoring approach to evaluate training

outcomes, and studying the impact of contextual factors (i.e., the COVID-19 pandemic) on training participation and motivation.

TABLE OF CONTENTS

	Page
TITLE PAGE	i
ABSTRACT	ii
LIST OF TABLES	vi
INTRODUCTION	1
Defining Diversity and Inclusion	2
Outcomes of Diversity	6
Diversity Management and Climate for Inclusion	10
Diversity Training as a Route to an Inclusive Workplace Culture	14
Diversity Training Approaches and Methodology	15
Diversity Training Outcomes	17
Present Study	20
Hypotheses	21
Additional Research Questions	24
METHOD	27
Participants	27
Procedure	28
Materials	29
RESULTS	32
Discussion of Analyses	32
A Brief Explanation of Propensity Scoring	33
Propensity Scoring Approach for the Present Study	37
Correlational Results	38
Hypothesis Testing Results	41
Research Question Analyses Results	43
DISCUSSION	48
Discussion of Hypothesis Testing and Results.....	48
Contributions to Diversity Training Research and Practice	51
Contributions to Research on the Impacts of COVID-19	53
Limitations and Directions for Future Research	56
Summary of Research Contributions and Practical Implications	59

Table of Contents (Continued)

	Page
CONCLUSION	60
REFERENCES	61
APPENDICES	68
Appendix A	69
Appendix B	71
Appendix C	72
Appendix D	74
Appendix E	75
Appendix F	76
Appendix G	77
Appendix H	78
Appendix I	79
Appendix J	80

LIST OF TABLES

Table	Page
Table 1. Summary of Training Programs Included in the Study	21
Table 2. Intercorrelations Among Study Variables Concerning the Focal Training Programs	39
Table 3. Intercorrelations Among Study Variables Concerning the Impact of COVID-19	40

INTRODUCTION:

THE EFFECTIVENESS OF A DIVERSITY AND INCLUSION INTERVENTION IN HIGHER EDUCATION IN THE CONTEXT OF COVID-19

The Society for Industrial and Organizational Psychology (SIOP) named Diversity, Inclusion, and Equity as number two on their list of Top 10 Workplace Trends of 2020 (Haynes, 2020). Similarly, the Society for Human Resource Management (SHRM) lists Diversity, Equity, and Inclusion among currently trending human resources topics (SHRM, 2020). Moreover, census data trends over the past several decades suggest that the United States (U.S.) is becoming older, more racially and ethnically diverse, and populated by increasing numbers of immigrants (Vespa, Medina, & Armstrong, 2020). By 2030, one fifth of Americans are projected to be over 65 years old, and by 2034, it is projected that older adults will outnumber children for the first time in U.S. history. It is also projected that by 2030, immigration will overtake natural increase (i.e., more American births than deaths) as the primary driver of U.S. population growth (Vespa et al., 2020), and by 2060, approximately one fifth of the American population will likely be foreign-born (Colby & Ortman, 2015). Additionally, by 2044, it is likely that more than half of all Americans will belong to a minority racial-ethnic group (i.e., any group besides non-Hispanic White alone) (Colby & Ortman, 2015).

These population projections have important implications for organizational policies designed to prevent discrimination based on age, race or ethnicity, and national origin in the workplace. As the working population becomes more diverse, it will become

increasingly essential that organizations are able to effectively utilize this diverse workforce. Consequentially, organizations must take steps to ensure equity in hiring, promotion, and termination and to create positive climates for inclusion in the workplace. However, despite the importance of diversity, equity, and inclusion, many organizations still fail to effectively implement these initiatives in the workplace. As such, recent literature has elaborated on both the benefits and pitfalls of diversity and how organizations can reap the benefits of diversity while avoiding the pitfalls by creating inclusive organizational climates. However, to fully grasp the impact of diversity and its implementation in the workplace, we must first understand what is meant by “diversity” and “inclusion.”

Defining Diversity and Inclusion

Diversity. In an organizational context, *diversity* generally refers to characteristics that differ between individuals. Diversity can be either more visible or less visible and can be more related or less related to job performance (Nishii & Mayer, 2009). Demographic characteristics, like age, race/ethnicity, and gender, are examples of diversity that is highly visible but unrelated to job performance. Experiential characteristics, like job tenure and education, are examples of diversity that is less visible but more related to job performance.

The concept of diversity has its roots in social identity theory. Tajfel (1982) defines social identity as “that part of the individuals’ self-concept which derives from their knowledge of their membership of a social group (or groups) together with the value and social significance attached to that membership” (p. 2). In other words, people define

and identify themselves through the groups to which they belong, and these groups have meaning within the larger social structure. As such, group identities affect the ways in which people from different groups interact with each other, and people prefer to be associated with groups that enjoy higher status and positive identities. Additionally, the similarity-attraction paradigm states that similarity between people in attitudes, values or demographic characteristics increases interpersonal attraction and liking (Brouwer & Boros, 2010). Therefore, high-status groups tend to accept and include people perceived to be similar to their group members, and they reject and exclude those who are different from the group members (Tajfel, 1982). This creates an in-group and an out-group, where in-group members are evaluated positively and out-group members are discriminated against (Brouwer & Boros, 2010). This boosts the self-esteem and psychological well-being of the in-group. In the context of organizational diversity, social identity theory suggests that people from diverse groups seek to identify with others in the organization who are like them (i.e., find their in-group), and people feel more included when the organization is welcoming and accepting of those who share their personal characteristics (Mor Barak et al., 2016).

The human tendency to associate with those like us and discriminate against those different from us can make it more difficult for organizations to effectively manage workplace diversity (Brouwer & Boros, 2010). This is problematic because diversity in organizations can be a competitive advantage. For example, Herring (2009) found that racial diversity in organizations was associated with increased sales revenue, more customers, greater market share, and greater relative profits, and gender diversity was

related to increased sales revenue, more customers, and greater relative profits. However, diversity must be properly managed for organizations to realize its benefits (O’Leary & Weathington, 2006). One major stumbling block for organizations attempting to unlock the benefits of diversity is that promoting diversity and multiculturalism can be perceived as a threat to the majority (or in-group) (Brouwer & Boros, 2010). As such, the idea of promoting diversity is better received if it is in line with the values, behaviors, and attitudes of the majority group; this allows diversity to become more of a competitive advantage than a stumbling block. Fortunately, organizations do have some power to change the values, behaviors, and attitudes of their workforce to be more accepting of diversity through the promotion of workplace inclusion. For example, the optimistic *contact hypothesis* suggests that contact in the workplace could give different groups the opportunity to find common ground and establish more positive relationships with one another. In other words, the contact hypothesis proposes that when the in-group and the out-group interact more often, more positive inter-group relations are established. However, positive inter-group relations are more likely to be established in situations where there are supportive egalitarian norms, common goals, equal status, cooperation, and voluntary opportunities for interaction. Ideally, in the workplace, these inter-group interactions eventually lead to everyone being considered part of the “in-group”, which brings us to the concept of inclusion.

Inclusion. In a workplace context, an inclusive organization typically refers to one in which “individuals of all backgrounds ... are fairly treated, valued for who they are, and included in core decision making” (Nishii, 2011, p. 1754). Additionally,

inclusive workplaces are committed to integrating diverse cultural identities, and they view these disparate identities as sources of insight and skill (Nishii, 2011). However, it can be a challenge for organizations to simultaneously achieve both diversity and inclusion, and, moreover, it is important to note that inclusive organizations do not want or expect workers with non-dominant cultural identities to assimilate to the values, attitudes, and behaviors of the cultural majority. Inclusion ties to the contact hypothesis in that the goal of an inclusive workplace is to promote positive inter-group relations. Additionally, the contact hypothesis' proposed situations that promote the development of positive inter-group relations (e.g., egalitarian norms, common goals, equal status, cooperation, and voluntary interaction) relate to some of the main tenets of an inclusive workplace. However, inclusion can also be tied to the optimal distinctiveness theory, which posits that people want to be like those around them, but they also want to feel accepted and appreciated for their unique characteristics and contributions (Mor Barak et al., 2016). In other words, people want to maintain and develop their own unique identities and cultures, while also continuing to fully belong to and participate in the greater community (Cunningham, 2015). Being included and accepted as part of a larger group and being able to express a unique individual identity are both very important to psychological health and well-being.

In practice, organizations seeking to develop a competitive edge with their diversity and inclusion initiatives should plan to (a) increase the number of groups represented within their workforce (i.e., increase diversity and representation) and (b) seek to make individuals from diverse groups feel appreciated for their unique

contributions (i.e., increase climate for inclusion). Taking both steps will help organizations benefit from the advantages of a diverse workforce while properly managing diversity. Creating an inclusive work environment ensures that each employee feels included and accepted in the organization and appreciated for his or her unique contributions.

Outcomes of Diversity

Over the past couple decades, researchers have attempted to determine what organizational outcomes are related to diversity in the workplace. Previous research has found that diversity can be related to positive outcomes, like commitment, satisfaction, and retention, but it can also be related to negative outcomes, like turnover, disengagement, and job stress (Mor Barak et al., 2016). However, it is important to note that different types of diversity may lead to different outcomes. For example, Nishii and Mayer (2009) distinguish between visible diversity, which includes attributes such as gender, race, and age, and invisible diversity, which includes attributes like education, organizational tenure, and socioeconomic status (SES). In this next section, I will discuss previous research findings about outcomes associated with various types of organizational diversity.

Education. Research on the impact of educational diversity within organizations has led to mixed results. However, some basic themes emerge. Previous research has found that those with higher educational status tend to be less satisfied and more likely to leave their organizations (Mor Barak et al., 2016). Additionally, those with specialized training for their jobs tend to have greater intentions to leave, but those with specialized

training also sometimes show greater job satisfaction. In consensus with some previous research, Mor Barak et al.'s meta-analysis found that lower education was associated with beneficial work outcomes, like job satisfaction and organizational commitment. Much of the relationship between educational level and job-related outcomes may be determined by labor market conditions and job opportunities, rather than individuals' affective reactions to their organizations. Nonetheless, based on previous research, organizations with a more educated and specialized workforce may expect higher turnover rates than organizations with less educated workforces. However, more positively, group-level education background variety is related to team creativity and innovation and to team performance, at top management levels (Bell et al., 2011).

Socioeconomic Status. Socioeconomic status (SES) diversity in the workplace can have varying effects on work outcomes, depending on the level of analysis. For example, at the individual level of analysis, people from lower SES backgrounds tend to have worse workplace outcomes than people from higher SES backgrounds, in that they are less likely to be hired and less likely to be promoted to leadership positions (Preston & De Graaf, 2019). However, at the group level of analysis, individuals from lower SES backgrounds tend to engage in more prosocial behaviors, and prosocial behavior leads to increased team cooperation and team performance. As such, teams that include more individuals from lower SES backgrounds may have better workplace outcomes.

Organizational Tenure. Findings about tenure diversity within organizations have been mixed. Although some studies have found longer tenure to be associated with lower burnout and higher organizational commitment, other studies have found the

opposite (Mor Barak et al., 2016). In indirect support of the former, Mor Barak et al.'s meta-analysis found that shorter organizational tenure was negatively associated with beneficial work outcomes. Much of the relationship between job tenure and organizational outcomes may be dependent upon other factors (e.g., emotional labor) that predict burnout and emotional exhaustion in certain positions, and age is also positively associated with tenure (Kunze, Boehm, & Bruch, 2013). As such, younger workers with shorter tenure may be experiencing a lack of positive outcomes due to their age rather than tenure. At the organizational level, improving employee tenure, along with its usual organizational benefits, may also improve individual-level employee outcomes.

Race and Ethnicity. Mor Barak and colleagues (2016) report on previous research that has found negative job-related outcomes for employees of racial and ethnic minorities. Affective job outcomes, like organizational commitment and job engagement, are especially negatively related to racial and ethnic minority status. For example, turnover intentions are positively related to racial and ethnic minority status, but minority status does not predict actual turnover very well. This is likely because job market conditions are often unfavorable to minorities. Mor Barak and colleagues (2016) suggest that these negative job outcomes are primarily the result of poor diversity management. Because racial and ethnic minorities are likely exposed to greater conflict and poorer treatment in the workplace, their affective reactions to the organization (e.g., commitment, engagement, turnover intentions) are negatively impacted. Due to these relationships between minority status and job-related outcomes, group-level organizational outcomes may be negatively impacted by racial and ethnic diversity in the

workforce, if there is poor diversity management within the organization or a low climate for inclusion. However, other studies have found more positive outcomes for racial diversity. For example, a study of organizations in the 1990s found that group-level racial diversity was associated with increased sales revenue, more customers, greater market share, and greater relative profits (Herring, 2009). These contradictory findings emphasize the importance of diversity management systems in the impact of racial diversity in the workplace.

Age. Age has a curvilinear relationship with organizational outcomes, suggesting that younger and older employees tend to experience the most negative outcomes (Mor Barak et al, 2016). For example, Zhang, Punnett, and Gore (2014) found that nurses younger than 40 and older than 60 reported greater intentions to leave. Similarly, in their meta-analysis, Mor Barak et al. (2016) found that younger workers were less likely to experience beneficial work outcomes, like job satisfaction, intentions to stay, and organizational commitment. As such, organizational-level outcomes may be negatively impacted if the workforce is either very young or very old and this type of diversity is poorly managed.

Gender. Previous studies on gender diversity have found mixed results. For example, Giffords (2009) found that a greater representation of women among workers is related to increased organizational support. Additionally, Herring (2009) found that organization-level gender diversity was associated with increased sales revenue, more customers, and greater relative profits. On the other hand, Bell, Villado, Lukasik, Belau, and Briggs (2011) found a negative relationship between team gender diversity and team

performance, and Mor Barak et al. (2016) found that being a man in a majority-female field was negatively associated with beneficial organizational outcomes. However, Nishii (2013) found that unit gender diversity was positively related to unit-level job satisfaction and justice climate and negatively related to turnover. Moreover, when a positive climate for inclusion was present, unit gender diversity was also negatively related to unit-level relationship conflict and task conflict.

In summary, previous research has found that group-level diversity and individual-level variation in key demographic and experiential traits can lead to both positive and negative individual and organizational outcomes. Beneficial outcomes of diversity can include job satisfaction, organizational and professional commitment, job tenure, and retention, while detrimental outcomes of diversity can include turnover, absenteeism, intentions to leave, job stress, emotional exhaustion, and depersonalization (Mor Barak et al., 2016). These beneficial and detrimental outcomes appear to be at odds, but organizational diversity management efforts towards creating a climate for inclusion can strongly impact the outcomes of workplace diversity.

Diversity Management and Climate for Inclusion

Part of effective diversity management is creating a climate for inclusion, which can increase the positive outcomes of diversity while decreasing the negative outcomes. This is well-demonstrated in a seminal study on inclusion climate by Nishii (2013). Nishii found that, within inclusive climates, interpersonal bias was reduced such that group-level gender diversity was related to lower levels of conflict. Moreover, she found that when work climate was inclusive, the negative effect that group conflict typically has

on unit-level satisfaction disappeared. As such, Nishii suggests that inclusion climate may be beneficial in reducing diversity-related turnover through increases in job satisfaction and reduced group conflict.

Similarly, Li, Lin, Tien, and Chen (2017) conducted a multi-level study with time-lagged data from 384 employees in 57 multicultural work teams, and they found that team cultural diversity was positively related to team creativity through team-level information sharing. They also found that team cultural diversity was positively related to individual-level creativity through employee-level information collaboration. Moreover, Li and colleagues found that the indirect positive relationship between team cultural diversity and team- and individual-level creativity was stronger when the climate for inclusion was stronger. They suggest that a stronger climate for inclusion might make team members more willing to share information with the rest of the team, thereby engendering creativity at the team and individual levels.

To further explore the beneficial outcomes of diversity and inclusion, Paolillo, Silva, and Pasini (2016) conducted a study on diversity climate and inclusion climate among Italian manufacturing organizations. They conceptualized diversity climate as “the degree to which a firm advocates fair human resource policy and socially integrates under-represented employees” (p. 310), whereas inclusion climate was conceptualized as “the degree to which individuals feel [they are] a part of critical organizational processes” (p. 311). Paolillo and colleagues found that diversity climate and inclusion climate were both positively related to safety participation motivation, and inclusion climate was also positively related to safety participation behavior. Moreover, they found that safety

participation motivation fully mediated the relationship between diversity climate and safety participation behavior, and it partially mediated the relationship between climate for inclusion and safety participation behaviors. This study emphasizes the tangible effects that a positive climate for diversity and inclusion can have on the workplace.

In a study on ability diversity among supervisors and subordinates, Dwertmann and Boehm (2016) found that incongruence in disability status among leadership dyads was related to lower ratings of leader-member exchange (LMX) and lower performance. LMX is the quality of the relationships that a leader develops with his or her followers (Nishii & Mayer, 2009). Dwertmann and Boehm (2016) found that LMX quality was worse among dyads where the supervisor had a disability than in dyads where the subordinate had a disability. However, they also found that unit-level inclusion climate acted as a buffer against the low LMX ratings typically found among dyads where the supervisor had a disability, but the subordinate did not. Dwertmann and Boehm's study emphasizes how important inclusion climate is for the management of all types of diversity.

Taking a slightly different angle, Konrad, Yang, and Maurer (2016) compared the efficacy of different approaches to creating an inclusive climate. They distinguish between classical disparity diversity and equality management systems (DEMS), institutional DEMS, and configurational DEMS. Classical disparity DEMS show limited development of any diversity management practices; institutional DEMS involve the instatement of specific selection systems and the monitoring of employee statistics, and configurational DEMS link diversity to overall business strategy. Konrad and colleagues

found that configurational DEMS seemed to most positively predict organizations' employment of workers with disabilities and visible minority groups, and configurational DEMS also positively predicted return on assets, which is an outcome measure linked to how effectively a company uses human assets and other assets. Based on this evidence, Konrad et al. suggest that organizations link their diversity management strategies to their overall business goals and strategies, as this will lead to the most positive results in terms of both hiring practices and tangible business outcomes.

All this empirical evidence points to the importance of inclusion climate and proper diversity management. However, in the past, many organizations simply increased employee diversity by adding more individuals from different backgrounds to the workplace, with little regard to how these individuals would be treated once onboard (Nishii, 2013). Subsequent research has shown that simply increasing the representation of women, or minorities, or people with disabilities, or people from different cultural backgrounds, etc. is not universally beneficial to organizations (Nishii, 2013). Rather, it is more important and beneficial for organizations to create a climate for inclusion, where individuals from all backgrounds are fairly treated, valued for who they are, and are included in core decision making (Nishii, 2013). The key to an inclusive climate is to not expect employees from nondominant social groups to assimilate to the norms of the majority. This ties into the concept of workplace equity, which suggests that organizations should account for differences in workers' backgrounds, privileges, and advantages (Putnam-Walkerly & Russell, 2016). Thus, in the modern-day workplace, it is quickly becoming the responsibility of the organization to not only reduce bias and

discrimination in hiring practices, but also to create an organizational climate that encourages a reduction in the typical sources of discrimination toward people of non-dominant social groups.

Diversity Training as a Route to Inclusive Workplace Cultures

The importance of creating an inclusive workplace climate has been well-established by previous research. However, this still leaves the question of *how* organizations can create and maintain inclusive work environments. The first step to this process is to establish policies that encourage equitable and unbiased personnel practices, including policies for hiring, promotion, and termination (Mor Barak et al., 2016). However, this approach may only increase diversity without creating an inclusive climate, and the implementation of these policies can be difficult to sustain without a concomitant change in organizational climate (Nishii, 2013). As such, many organizations may turn to diversity and inclusion (D&I) training to fill the gap between policy and implementation and create a more inclusive climate. In fact, as many as two-thirds of U. S. organizations report using some sort of D&I training in the workplace (Kalinowski, Steele-Johnson, Peyton, Leas, Steinke, & Bowling, 2013).

The idea behind diversity training is to address prejudice, stereotyping, and other biases; it is supposed to facilitate positive intergroup interactions, reduce prejudice and discrimination, and enhance the knowledge, skills, and motivation of people to interact with diverse others (Bezrukova, Jehn, & Spell, 2012). However, researchers have been continually investigating how well these types of organizational interventions truly work. For example, Duguid and Thomas-Hunt (2015) found that making people more aware of

prevalent stereotypes sometimes exacerbates subsequent stereotypic attitudes about and behavior toward out-group members. Diversity training sometimes includes the “stereotype discrediting” method, which involves directly confronting common stereotypes about a marginalized group (Lindsey, King, Hebl, & Levine, 2015). However, based on Duguid and Thomas-Hunt’s (2015) findings, this strategy may be counteractive to promoting diversity and acceptance and improving organizational diversity management systems. These findings highlight the importance of thorough research in implementing the most effective methods for creating inclusive climates in organizations.

Diversity Training Approaches and Methodology

There are a variety of methods that can be implemented when administering D&I training, and research has found some methods to be more effective than others. In their review of diversity training research, Bezrukova and colleagues (2012) distinguish between certain design features often present in diversity trainings. For example, the focus of diversity training material is often either on group-specific topics (e.g., race, gender) or it emphasizes inclusiveness across multiple groups. In their 2016 meta-analysis, Bezrukova, Spell, Perry, and Jehn found that there was no significant difference between the effectiveness of these two diversity training approaches. As such, organizations should feel free to tailor their D&I training programs to be either more inclusive or more group-specific based on which approach they think will be best received by the trainees they are targeting.

Diversity training can also be split according to the types of outcomes being targeted, namely awareness, behavior, or a combination of both (Bezrukova et al., 2012; 2016). Awareness training promotes trainees' self-awareness on diversity-related topics, like cognitive biases and heuristics that affect the interpretation of others' behavior. Behavioral training, on the other hand, educates trainees' on how to monitor their own actions to appropriately respond to different situations in the workplace (e.g., discouraging race- or gender-based jokes). It should be noted that behavioral training is often conducted in conjunction with awareness training. In their 2016 meta-analysis, Bezrukova et al. found that behavior-based training in combination with awareness training and behavior-based training alone both had more impact on affective and behavioral learning outcomes than awareness training alone. As such, organizations should incorporate a behavioral component into their D&I training programs when possible.

Although sexual minority status is one area of workplace diversity that has been studied less often than other types of diversity, there is some evidence that this particular type of diversity is well-managed with diversity training (Rosopa, Fynes, D'Souza, & Xoxakos, 2020). For example, Lindsey and colleagues (2015) conducted a study on diversity training programs promoting lesbian, gay, and bisexual (LGB) inclusion, elaborating on some different ways diversity training material can be framed beyond Bezrukova et al.'s (2012) distinctions. Lindsey and colleagues (2015) compared three different interactive diversity training approaches, including perspective taking, goal setting, and stereotype discrediting. The perspective taking approach is designed to

reduce prejudice by requiring individuals to imagine what it would be like to be part of a stigmatized group. This breaks down in-group bias and reduces “us versus them” mentality. The goal setting diversity training strategy involves trainees setting challenging, but achievable, diversity-related goals involving a specific stigmatized group. Last, the stereotype discrediting strategy, as previously mentioned, involves asking trainees to engage in actively discrediting common stereotypes about a stigmatized group.

Lindsey et al. (2015) found that perspective taking seemed to be the most effective training method for increasing supportive behaviors toward LGB individuals after a time lapse of eight months. The authors suggest that perspective taking may increase trainees’ internal motivation to respond without prejudice, while the other two approaches may only increase external motivation to respond without prejudice. However, a similar study by Madera, King, and Hebl (2013) successfully used the goal setting approach to improve attitudes toward sexual minorities. These research results emphasize that there are often multiple routes to success in diversity training.

Diversity Training Outcomes

The success of diversity training, and training in general, can be evaluated using several different methods, including trainee reactions, affective outcomes, behavioral outcomes, cognitive and learning outcomes, and results criteria. Reactions to training refer to trainees’ subjective evaluations about the learning experience and are typically measured via self-report questionnaires (Arthur, Bennett, Edens, & Bell, 2003; Sitzmann, Brown, Casper, Ely, & Zimmerman, 2008). Affective outcomes include changes in

attitudes, self-efficacy, and motivation in general (Kalinowski et al., 2013). Behavioral outcomes include changes in trainees' job-related behaviors or performance, and learning or cognitive outcomes seek to determine what knowledge participants learned from their training experiences (Arthur et al., 2003). Last, results criteria typically attempt to measure the utility of the training program to the administering organization (e.g., in terms of productivity, company profits, etc.).

In their meta-analysis on the outcomes of diversity training, Bezrukova and colleagues (2016) found an overall effect size (Hedges g) for diversity training of 0.38. Additionally, diversity training seemed to have the largest effect on reactions to training (.61) and cognitive learning (.57), while smaller effects were found for behavioral (.48) and affective learning (.30). These results are somewhat consistent with previous meta-analytic findings by Arthur and colleagues (2003), who found that evaluations using trainee reactions yielded more consistently positive effects than learning, behavioral, or results criteria. However, Arthur et al. also found that learning criteria tended to yield the largest effect sizes overall for training programs. Bezrukova and colleagues (2016) additionally found that diversity training programs tended to be most effective when they were supplemented by other diversity initiatives targeting awareness and skills development.

Kalinowski and colleagues (2013) also conducted a meta-analysis on diversity training outcomes, finding an overall effect size (Cohen's d) for diversity training of .39, which is very consistent with results from Bezrukova et al. (2016). Kalinowski et al. (2013) also compared effect sizes based on evaluation methods, and they found that knowledge

and cognitive outcomes seemed to have the largest effect size (.62), followed by on-the-job behavioral outcomes (.35), then affective outcomes (.27), and attitudes seemed to be least affected (.23). These results are also very consistent with the findings of Bezrukova and colleagues (2016). This research evidence suggests that diversity training can have a positive, meaningful impact on trainees. However, diversity training seems to have a larger impact on cognitive and behavioral outcomes than on affective and attitude outcomes. As such, Kalinoski and colleagues (2013) suggest that future diversity training efforts consider alternative ways to measure attitudes. For example, attitudes are often only measured post-training, but that does not account for the possibility that attitudes were favorable before training, so pre- and post-training within-subjects measurement should be used more often. Additionally, only explicit attitudes are often assessed, but comparing implicit attitudes pre- and post-training might reveal more attitude change than comparing explicit attitudes would. Even with these changes, however, Kalinoski and colleagues suggest that attitudes may simply be more difficult to change than behavior or knowledge.

A study by Carnes and colleagues (2015) attempted to address one of the attitude outcome evaluation issues raised by Kalinoski and colleagues (2013) by assessing both implicit and explicit attitude changes after their diversity training intervention about gender bias. Carnes and colleagues' (2015) diversity intervention successfully increased awareness of personal gender bias, motivation to promote gender equality, self-efficacy for enacting gender equality, and expectations of positive gender equity outcomes among participants. However, the intervention was unsuccessful in changing implicit biases.

These research results indicate that Kalinoski and colleagues (2013) may be correct in suggesting that attitudes, especially implicit attitudes, may simply be more difficult to change than behavioral and cognitive outcomes. However, evidence from both individual studies (e.g., Carnes et al., 2015) and meta-analyses (e.g., Bezrukova et al., 2016; Kalinoski et al., 2013) indicates that, while the effect sizes are smaller compared to other outcomes, attitudes can be positively impacted by diversity training programs.

Present Study

The current study was part of an initiative to create a more inclusive environment for women and minorities among the STEM faculty at a large, southeastern university. As part of this initiative, over several years, the university offered several diversity training programs designed to help participants recognize bias in the workplace, teach participants methods they can use to confront bias, and help participants prevent those biases from impacting work decisions (e.g., hiring, promotion, leadership selection, etc.). See Table 1 for a summary of the different training programs offered as part of this initiative that were included in the present study.

Table 1. Summary of training programs included in the study

Program Name	Program Learning Goals	Outcome Goals	Curriculum Focus
TIGER Advocates	Recognizing covert bias and departmental bias, along with monitoring institutional climate.	Active promotion of gender diversity and equality. Advocating for the hiring and promotion of women and other underrepresented faculty.	Awareness and behavior
Towards Equitable Workloads	Implicit bias awareness and reduction, identifying the pervasiveness of implicit bias in the workplace, and learning strategies to mitigate the effects of bias.	Increasing the fairness of workload distributions, with a focus on avoiding allocating assignments based on gender stereotypes.	Awareness and behavior
Tenure, Promotion, and Reappointment (TPR) Equitable Evaluation Training	Awareness of implicit and explicit racial and gender biases and how those biases can impact decisions about faculty careers. Identification of the differential impacts of COVID-19 on women faculty and faculty with young children.	Promoting more equitable and unbiased practices in TPR decisions and performance evaluations, with an emphasis on awareness of gender- and race-specific challenges.	Awareness
Trailblazers Leadership Program	Providing cohort-based leadership training to prepare faculty for leadership roles while furthering institutional diversity.	Promoting individual development, networking, mentoring, and advancing gender equality at the institution.	Awareness and behavior

Hypotheses

Previous research has found that diversity training can have a tangible impact on explicit attitudes (e.g., Bezrukova et al., 2016; Kalinoski et al., 2013). Additionally, there is evidence from previous research that diversity training can promote attitudes favoring

gender equality (e.g., Carnes et al., 2015) and can decrease racist attitudes (e.g., Devine et al., 2012). As such, I proposed the following:

Hypothesis 1. Compared to those who completed no training, participants who completed any training would have (a) greater concern about gender discrimination, (b) greater concern about racial discrimination, (c) greater perceptions of inclusion climate, (d) more positive attitudes towards workplace diversity, (e) lower levels of modern sexism, and (f) lower levels of modern racism.

However, it should be noted that the different training programs offered by the university all had slightly different target participants and goals. As such, we can expect different trainings to impact different attitudinal outcomes. The different trainings and the outcomes they targeted are listed in more detail below.

TIGER Advocates. TIGER Advocates taught participants to recognize micro- and macro-level biases against women and minorities and taught them how to intervene when these biases occur. Trainees were mentored on issues like covert bias, institutional bias, and institutional climate during a two-hour session. After they completed their training, participants were expected to be active proponents of diversity and equality at the university, and they were encouraged to intervene when they could increase the representation of women and minority faculty through hiring and promotion decisions. Additionally, trainees were encouraged to promote the fair and equitable treatment of all faculty within their academic units. As such, I expected:

Hypothesis 2. Compared to the control group, trainees who completed the TIGER Advocates training were expected to display (a) lower modern sexism, (b) lower modern racism, (c) greater concern about gender discrimination, (d) greater concern about racial discrimination, (e) greater perceptions of inclusion climate, and (f) more positive attitudes toward workplace diversity.

Equitable Workloads Training. This training program was targeted at department chairs, and its goal was to increase awareness of implicit gender bias in department workload allocations. During short workshops, trainees were taught about implicit gender bias and were taught strategies to mitigate bias through accurate time use tracking, equitable workload allocations, and fair performance reviews. After the training, department heads were expected to promote more gender-equitable practices within their academic units. As such, I expected:

Hypothesis 3. Compared to the control group, Equitable Workloads trainees would display (a) greater concern about gender discrimination, (b) lower modern sexism, (c) greater perceptions of inclusion climate, and (d) more positive attitudes toward workplace diversity.

Tenure, Promotion, and Reappointment (TPR) Equitable Evaluation

Training. TPR committee chairs were the target group for this intervention. This training was designed to promote more equitable practices in TPR decisions and performance evaluations by making TPR committee chairs more aware of implicit and explicit racial and gender biases and how those biases can impact important decisions about faculty

careers. Additionally, this training took place during the pandemic caused by the 2019 novel coronavirus (COVID-19), so the differential impacts of COVID-19 on women faculty and faculty with young children were also discussed. As such, I expected:

Hypothesis 4. Compared to the control group, trainees who completed the TPR training were expected to display (a) lower modern sexism, (b) lower modern racism, (c) higher concern about gender discrimination, (d) higher concern about racial discrimination, (e) greater perceptions of inclusion climate, and (f) more positive attitudes toward workplace diversity.

Trailblazers. The Trailblazers program was a leadership initiative focusing on gender equality in the STEM fields. This training prepared faculty for leadership roles in either academic or professional environments by having trainees participate in a nine-month long program. During the program, participants received leadership training, individual development opportunities, experience working on projects promoting gender equality, mentorship, and networking opportunities.

Hypothesis 5. Compared to the control group, those who completed the Trailblazers program, would display (a) greater concern about gender discrimination, (b) lower modern sexism, (c) higher perceptions of inclusion climate, and (d) more positive attitudes toward workplace diversity.

Additional Research Questions

Data for the present study were collected during the unique challenges caused by COVID-19. Additionally, although most of the training interventions began before the

pandemic, many were continued during COVID-19, and some trainings (e.g., TPR Equitable Evaluation) were initiated after the pandemic had begun. As such, it was essential to consider the potential impact of COVID-19 on the results of the training interventions.

COVID-19 has substantially impacted the way people work in the U.S. Reports starting in April 2020 showed that nearly half of all currently employed U. S. workers were working remotely (Brynjolfsson, Horton, Ozimek, Rock, Sharma, & Ye, 2020). Of employed workers in April 2020, over 34% had switched from commuting to remote work due to the pandemic. In the present study, faculty at the university switched to remote teaching in mid-March 2020, and they had the option to resume in-person teaching in late September 2020. However, only about 20% of classes at the university had an in-person component in fall 2020. The option of in-person teaching continued at the university until late November 2020, after which all faculty switched to completely remote teaching again until the start of the spring 2021 semester. Students at the university were also able to request fully online learning for the fall semester, and approximately 15% of the student population chose to do this. Data collection for the present study began in March 2021. Approximately half of the courses offered at the university for the spring 2021 semester included at least some in-person teaching. In summary, the faculty targeted by the present study switched between working remotely and working in-person for the duration of the pandemic.

This influx of remote work, paired with increased home production demands due to the closures of schools and daycares, may have impacted faculty productivity during

the pandemic (Brynjolfsson et al., 2020). Preliminary research by Myers and colleagues (2020) showed that academic researchers with partners and with young dependents experienced the most drastic declines in time dedicated to research since the start of the pandemic. As such, this study sought to answer the following research questions:

Research Question 1. Is there a relationship between childcare responsibilities and faculty willingness and ability to participate in voluntary learning opportunities, like voluntary trainings, diversity and inclusion initiatives, and continuing education seminars?

Research Question 2. Is there a relationship household responsibilities and faculty willingness and ability to participate in voluntary learning opportunities, like voluntary trainings, diversity and inclusion initiatives, and continuing education seminars?

Preliminary research also showed that many researchers were dedicating less time to work, including researching, fundraising, teaching, and other work tasks, after the start of the pandemic (Myers et al., 2020). Thus, this study sought to answer the following research questions:

Research Question 3. Is there a relationship between changes in summer research hours during COVID-19 and faculty willingness and ability to participate in voluntary learning opportunities, like voluntary trainings, diversity and inclusion initiatives, and continuing education seminars?

Research Question 4. Is there a relationship between changes in academic year research hours during COVID-19 and faculty willingness

and ability to participate in voluntary learning opportunities, like voluntary trainings, diversity and inclusion initiatives, and continuing education seminars?

Research Question 5. Is there a relationship between changes in summer teaching hours during COVID-19 and faculty willingness and ability to participate in voluntary learning opportunities, like voluntary trainings, diversity and inclusion initiatives, and continuing education seminars?

Research Question 6. Is there a relationship between changes in academic year teaching hours during COVID-19 and faculty willingness and ability to participate in voluntary learning opportunities, like voluntary trainings, diversity and inclusion initiatives, and continuing education seminars?

METHOD

Participants

Participants in the present study included 79 faculty at a large, southeastern university. The sample included tenured (58.2%), tenure-track (32.9%) and non-tenure-track faculty (8.9%), men (44.3%) and women (55.7%), and multiple races and ethnicities (84.8% white, 15.2% people of color). All eight university colleges were represented within the sample, and faculty from a variety of departments were included as well. Additionally, participants from all of the focal training programs were represented in the sample. About 38% of the sample had participated in at least one of the focal trainings. As for the specific trainings, 20.3% had participated in TIGER

Advocates, 19.2% had participated in Trailblazers, 10.1% had participated in TPR Committee Equitable Faculty Review, and 5.1% had participated in Towards Equitable Workloads.

Procedure

Data for the present study were collected online via a survey administered to a sample of 620 faculty at the institution. An initial email survey invitation was sent to the sample, and then a reminder email was sent to the same sample about two weeks later. There was a 12.7% response rate. The survey was administered entirely online via Qualtrics survey software, and all survey links were anonymous. Once they clicked on the link, faculty read the informed consent and chose to participate in the survey or not. If they chose to continue, faculty followed instructions to enter a string of numbers and letters that was used as a unique identifier. Participants then completed a demographics questionnaire, followed by a training checklist, on which they indicated which education programs, trainings, and interventions they had completed while employed at the university. Then, participants completed several survey blocks of measures, including a COVID-19 impact questionnaire, a modern sexism scale, a concern about gender discrimination measure, a concern about racial discrimination measure, a modern racism measure, a climate for inclusion scale, and an attitude toward diversity scale. At the end of the survey, participants were given the opportunity to provide their email address to enter a drawing for one of 10 gift cards worth \$20 each. Participants were informed of this opportunity both in the informed consent and in the instructions for the email question.

Materials

Demographics Questionnaire. Participants completed a demographics questionnaire, indicating their gender identification, birth year, race and/or ethnicity, college and department, academic rank, and tenure status. Please refer to Appendix A for this questionnaire.

Training Checklist. Participants viewed a checklist with the following items: TIGER Advocates; Towards Equitable Workloads Training; Tenure, Promotion, and Reappointment (TPR) Committee Equitable Faculty Review Training; Trailblazers Leadership Training, or none of the above. Participants will be asked to check the boxes for the trainings and interventions they have participated in while employed at the university. Participants were then asked to indicate what year they had participated in each training or intervention they checked off. Please refer to Appendix B for this checklist.

COVID-19 Impact Questionnaire. The present study was conducted during the global COVID-19 pandemic, which substantially impacted people's lives. Therefore, we included some exploratory research questions to determine what aspects of life and work during COVID-19 may have impacted faculty decisions to participate in non-mandatory events, like diversity and inclusion initiatives. As such, participants were asked a series of questions designed to assess the impact COVID-19 had on their teaching and research hours and productivity, and its impact on their willingness and ability to participate in voluntary training, diversity and inclusion initiatives, and continuing education seminars.

Additionally, participants were asked if childcare and household responsibilities impacted their productivity during COVID-19. Please refer to Appendix C for this questionnaire.

Attitudinal Measures. After answering the demographics questionnaire, training checklist, and exploratory COVID-19 research questions, participants were prompted to respond to a series of attitudinal questionnaires, listed next.

Modern Sexism. The Swim, Aikin, Hall, and Hunter (1995) modern sexism scale includes eight items divided into three subcategories: denial of continuing discrimination, antagonism towards women's demands, and resentment about special favors for women. Participants rate each statement on a seven-point response scale (1 = strongly agree, 7 = strongly disagree). A sample item is "Discrimination against women is no longer a problem in the United States." Higher scores indicate greater endorsement of modern sexist statements. For the current sample, $M = 2.05$, $SD = 1.01$, $\alpha = .804$. Please refer to Appendix D for this scale.

Concern about Gender Discrimination. The concern about gender discrimination scale is based on the concern about racial discrimination scale from Devine, Forscher, Austin, and Cox (2012). This scale consists of four items, which were modified for the present study to measure concern about gender discrimination rather than racial discrimination. Items are rated on a 10-point response scale (1 = strongly disagree, 10 = strongly agree). A sample item is "I consider gender discrimination to be a serious social problem." Higher scores on this scale indicate greater concern about gender discrimination. For the current sample, $M = 8.61$, $SD = 1.80$, $\alpha = .624$. Please refer to Appendix E for this scale.

Concern about Racial Discrimination. The original concern about racial discrimination scale by Devine and colleagues (2012) was also included. This scale consists of four items scored on a 10-point response scale (1 = strongly disagree, 10 = strongly agree). Higher scores indicate greater concern about racial discrimination. A sample item is “I consider racial discrimination to be a serious social problem.” For the current sample, $M = 9.38$, $SD = 1.49$, $\alpha = .833$. Please refer to Appendix F for this scale.

Modern Racism. This scale, taken from Swim and colleagues (1995), includes six items rated on a seven-point response scale (1= strongly disagree, 7 = strongly agree). The scale is divided into three subsections, including denial of continuing discrimination, antagonism toward African-Americans’ demands, and resentment about special favors for African-Americans. Higher scores indicate greater endorsement of racist statements. A sample item is “Over the past few years, Black people have gotten more economically than they deserve.” For the current sample, $M = 1.42$, $SD = .751$, $\alpha = .704$. Please refer to Appendix G for this scale.

Climate for Inclusion. Nishii’s (2013) short-form climate for inclusion scale includes 15 items scored on a five-point response scale (1 = strongly disagree, 5 = strongly agree). Higher scores indicate greater perceptions of a positive inclusion climate. Sample items include “Employees in this department receive ‘equal pay for equal work’” and “This department has a culture in which employees appreciate the differences that people bring to the workplace.” For the current sample, $M = 3.41$, $SD = .975$, $\alpha = .956$. Please refer to Appendix H for this scale.

Workplace Diversity Scale. This scale measures perceptions of diversity in the workplace (De Meuse & Hostager, 2001). It consists of 20 items rated on a five-point response scale (1 = disagree, 5 = agree). Higher scores indicate more positive perceptions of diversity at work. Sample items include “I believe diversity is fair” and “I participate in organizational diversity efforts.” For the current sample, with the scale including both positively worded items and reversed negatively worded items, $M = 4.43$, $SD = .436$, $\alpha = .856$ (positive items only: $M = 4.46$, $SD = .446$, $\alpha = .791$; negative items only: $M = 1.60$, $SD = .547$, $\alpha = .794$). Please refer to Appendix I for this scale.

RESULTS

Discussion of Analyses

In the current study, random assignment to the training and control groups was not possible because all the university training programs included in the study were voluntary trainings. As such, there may have been a selection bias present, in that faculty who chose to complete training may have systematically varied from faculty who chose not to complete the voluntary trainings. If this was the case, it would have introduced confounding variables into the study, which would have made it difficult or impossible to interpret whether any outcome differences between the training and control groups are due to the training itself or are merely a function of the baseline characteristics of each group. These issues of selection bias and confounding variables are part of the reason that randomized controlled trials are considered the gold standard for valid causal inferences (Austin, 2011). However, in cases where randomized controlled trials are not feasible, such as in this study, propensity scoring can be used to mimic some of the characteristics

of a random sample and random assignment to treatment conditions. Propensity scoring makes the training and control groups as similar as possible on all measured confounding variables so that comparisons between the two groups on outcome variables are based only on differences in training participation. This attenuates the impact of selection bias and confounds on the results of the study.

A Brief Explanation of Propensity Scoring

From a statistical perspective, a propensity score is “the probability of treatment assignment conditional on observed baseline characteristics” (Austin, 2011, p. 399). In practice, the propensity score balances out the distribution of observed baseline characteristics between the treatment and control groups so that they are more similar (Austin, 2011). In other words, propensity scores allow researchers to analyze a nonrandomized study as if it were a randomized controlled trial. Most commonly, a propensity score is estimated using a logistic regression model, in which treatment status is regressed on observed baseline characteristics (Luellan et al., 2005); in this case, the estimated propensity score is the predicted probability of treatment derived from the fitted regression model. It should be noted that other methods of deriving propensity scores exist, including bagging, neural networks, classification trees, and others (Austin, 2011; Luellan et al., 2005). However, when comparing the logistic regression method to the more complex methods of bagging and classification trees, Luellan and colleagues (2005) could not conclude which of the three methods resulted in the most accurate estimates of treatment effects, and no single model resulted in the greatest reduction in bias for all measured outcomes. Moreover, computationally complex methods of

computing propensity scores, like classification trees, often overfit the data, which must be adjusted for in subsequent analyses, and some of these methods are better suited for propensity score stratification than for propensity score matching. As such, I used the logistic regression method to generate propensity scores for simplicity and ease of interpretation. Additionally, there are several different approaches to selecting a control group using propensity scores, including: matching on the propensity score, stratification on the propensity score, inverse probability of treatment weighting using the propensity score, and covariate adjustment using the propensity score.

Propensity score matching is the most common approach to propensity scoring, and it involves forming matched sets of treated and untreated subjects who share a similar value for the propensity score (Austin, 2011). Usually, this is implemented using one-to-one pair matching. Once these matched pairs are formed, then the treatment effect can be easily estimated by comparing the outcomes between treated and untreated participants in the matched sample. For continuous outcomes (e.g., as in the current study), treatment effects are estimated as the difference between the mean outcome for the treated sample and the mean outcome for the control sample. Whenever this approach to propensity scoring is used, the lack of independence between the propensity score matched samples should be accounted for when estimating the variance of the treatment effects. As such, the present study would require the use of paired samples *t*-tests for hypothesis testing. Additionally, regarding the way matched pairs were created for the present study, I used matching without replacement to avoid having some untreated subjects paired with multiple treated subjects, and based on previous research recommendations, I used

nearest neighbor matching within a specified caliper distance (discussed in the next section titled *Propensity Scoring Approach for the Present Study*), which allowed me to constrain how far apart matched subject propensity scores could acceptably be.

An alternative to propensity score matching is propensity score stratification, which involves stratifying subjects into mutually exclusive subsets based on their estimated propensity score (Austin, 2011). In this scenario, subjects are ranked based on their estimated propensity score, and then they are stratified into subsets based on previously defined thresholds of propensity scores. Often, these thresholds are based on the quintiles of propensity scores, such that the subjects are all divided into five equally sized groups. Then, treatment effects are estimated within each of the five strata, and the strata can be pooled to estimate an overall treatment effect. Propensity score stratification may be less efficient than propensity score matching at eliminating differences between treatment and control groups (Austin, 2011), but it is also less sensitive to nonlinearity in the relationships between propensity scores and outcomes (Luellen et al., 2005).

Another alternative to propensity score matching is inverse probability of treatment weighting (Austin, 2011). This method uses weights based on the propensity score to create a synthetic sample in which the measured baseline attributes are independent of treatment assignment. Each subject's weight is equal to the inverse of the probability of receiving the treatment that the subject actually received. As such, regression models can be weighted by these inverse probabilities of treatment to more accurately estimate the causal effects of the treatment. However, with this method, the

variance estimation must account for the weighted nature of the synthetic sample, and often this includes using robust variance estimation.

The last alternative to propensity score matching is covariate adjustment using the propensity score (Austin, 2011). In this approach, the outcome variable is regressed on an indicator variable denoting treatment status and the estimated propensity score. For continuous outcomes (e.g., as in the present study), a linear regression is conducted, and the effect of treatment is determined by the estimated regression coefficient from the fitted regression model. When using a linear model, the treatment effect is an adjusted difference in means. However, this method assumes that the nature of the relationship between the propensity score and the treatment outcome has been correctly modeled.

Comparing the different approaches to propensity scoring, previous research has found that propensity score matching eliminates a greater proportion of systematic differences in baseline characteristics between the treatment and control groups than does propensity score stratification and covariate adjustment using the propensity score (Austin, 2011). Propensity score matching and inverse probability of treatment weighting based on the propensity score are more comparable in removing systematic differences between the treatment and control groups, but in some settings, propensity score matching outperforms inverse probability of treatment weighting as well. Based on this evidence, I used propensity score matching for my method of selecting control groups to compare to my training groups. All analyses were conducted in R (R Core Team, 2021), using the MatchIt package (Ho, Imai, King, & Stuart, 2011).

Propensity Scoring Approach for the Present Study

Based on previous research (e.g., Carnes et al., 2015), I expected that university college and faculty rank, gender, and race/ethnicity would impact which faculty attended trainings and which did not. As such, these attributes were entered as covariates in the logistic regression for the propensity score assignment process, and then control groups were matched to trainee groups using the propensity score matching approach (i.e., one-to-one matching of trainees to controls). This approach should have effectively eliminated the baseline differences between the training groups and the control groups which may otherwise have impacted the outcomes of the trainings and initiatives involved in the present study (Austin, 2011). It should be noted that there are multiple methods of forming matched control-trainee pairs, and, as such, the following section discusses the approach used in the present study.

First, matching without replacement was used rather than matching with replacement. This means that once a control subject was matched to a trainee subject, that subject could not be matched again to another trainee subject. However, the individual training programs were each treated as separate and independent analyses in this study. As such, if a participant took more than one training, he or she may have been used more than once as a trainee matched to a control in a separate analysis. Similarly, one control participant could have been used more than once, but only over separate and independent analyses.

Additionally, matches in this study were selected using the closest neighbor within a specified caliper distance approach. This approach selects the control subject

with the closest propensity score to a given trainee subject, if it is within a specified threshold. This constricts the maximum acceptable difference between the propensity scores of each match. Previous research suggests that the optimal caliper distance to use for this threshold is equal to 0.2 of the standard deviation of the logit of the propensity score; using this value, or one close to it, appears to minimize the mean squared error of the estimated treatment effect (Austin, 2011), so this standard was used in the present study.

Austin (2011) suggests that variance estimators that account for propensity score matching more accurately reflect the sampling variability of the estimated treatment effect. As such, it is suggested that paired t tests should be used when assessing the statistical significance of treatment effects on a continuous outcome, so I used paired t tests in my subsequent analyses. Once the control groups were selected for each of the trainee groups based on propensity score matching, hypothesis testing was conducted. The number of matched trainee-control pairs generated varied based on the training program being evaluated and the hypothesis being tested. The number of matched pairs on which the paired t tests were conducted is indicated in each hypothesis results section.

Correlational Results

Initial correlational analyses were conducted to investigate the general linear relationships between all study variables. See Tables 2 and 3.

Table 2. Intercorrelations among study variables concerning the focal training programs

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13
1 Gender	0.44	0.5													
2 Academic rank, title, or position	2.25	1.13	-.384**												
3 Tenure Status	1.51	0.658	-.223*	.880**											
4 TIGER Advocates	0.2	0.404	.311**	-.283*	-.342**										
5 Towards Equitable Workloads Training	0.05	0.221	0.143	-0.155	-0.091	0.171									
6 TPR Committee Equitable Evaluation Training	0.1	0.304	0.038	-0.151	-0.132	-0.169	.305**								
7 Trailblazers Leadership Program	0.13	0.335	-0.186	-0.154	-0.178	-0.097	-0.088	.251*							
8 Any Training at All	0.38	0.488	0.142	-.363**	-.407**	.644**	.295**	.429**	.487**						
9 Concern for Gender Discrimination	8.61	1.8	-.271*	0.046	-0.053	0.044	-0.095	-0.019	0.153	0.146					
10 Modern Sexism	2.05	1.01	.240*	-0.08	0.041	-0.154	0.128	.233*	-0.113	-0.091	-.745**				
11 Concern for Racial Discrimination	9.38	1.49	-.233*	0.015	-0.139	0.164	-0.158	-0.111	0.115	0.133	.784**	-.817**			
12 Modern Racism	1.42	0.751	.292*	-0.023	0.099	0.003	0.143	0.195	-0.047	0.063	-.637**	.730**	-.688**		
13 Perceptions of Inclusion Climate	3.41	0.975	0.08	-0.181	-0.085	.230*	0.09	0.18	0.165	.417**	0.063	-0.091	-0.029	0.061	
14 Attitudes toward Diversity	4.43	0.436	-0.14	-0.09	-0.162	0.1	-0.09	0.059	.227*	.259*	.515**	-.426**	.397**	-.299**	.367**

Notes. $N = 76$. Gender is coded as 0 = Female, 1 = Male. Academic Rank/Title/Position coded as 1 = Professor, 2 = Associate Professor, 3 = Assistant Professor, 4 = Instructor, 5 = Lecturer, 6 = Other. Tenure status coded as 1 = tenured, 2 = On tenure track but not tenured, 3 = Not on tenure track. ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

Table 3. Intercorrelations among study variables concerning the impact of COVID-19

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11
1 Gender	0.44	0.5											
2 Academic Rank/Title/Position	2.25	1.13	-.384**										
3 Tenure Status	1.51	0.658	-.223*	.880**									
4 Training Participation	0.38	0.488	0.142	-.363**	-.407**								
5 Summer Research Hours during COVID-19	2.41	1.2	0.039	0.018	0.029	0.04							
6 Academic Year Research Hours during COVID-19	2.22	1.18	0.053	-0.013	0.056	0.034	.766**						
7 Summer Teaching Hours during COVID-19	3.47	0.963	-0.049	0.005	0.067	0.015	-0.083	-0.036					
8 Academic Year Teaching Hours during COVID-19	4.17	0.932	-0.095	-0.008	0.019	-0.178	0.096	-0.037	.414**				
9 Significant Childcare Responsibilities	0.47	0.502	-0.071	0.131	0.03	0.212	-.325**	-0.153	0.073	0.041			
10 Significant Household Responsibilities	0.87	0.336	-0.035	0.055	0.004	.292**	0.102	0.138	0.153	-0.008	.362**		
11 Ability to participate in voluntary learning opportunities	2.35	1.14	-0.074	0.136	0.209	-0.063	.362**	.448**	0.108	0.006	-.266*	.114	
12 Ability to do DEI Initiatives during COVID-19	2.44	1.12	-0.065	0.165	0.221	-0.049	.384**	.456**	.238**	0.074	-0.161	0.107	.715**

Notes. *N* = 77. Gender is coded as 0 = Female, 1 = Male. Academic Rank/Title/Position coded as 1 = Professor, 2 = Associate Professor, 3 = Assistant Professor, 4 = Instructor, 5 = Lecturer, 6 = Other. Tenure status coded as 1 = tenured, 2 = On tenure track but not tenured, 3 = Not on tenure track. Research and teaching hours are coded as 1 = decreased significantly, 2 = decreased slightly, 3 = remained the same, 4 = increased slightly, 5 = increased significantly. Significant childcare and household responsibilities are coded as 0 = No, 1 = Yes. Ability/Willingness variables are coded as 1 = significantly less able/willing to participate, 2 = slightly less able/willing to participate, 3 = no impact on ability/willingness to participate, 4 = slightly more able/willing to participate, 5 = significantly more able/willing to participate. ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

Hypothesis Testing Results

Hypothesis 1 (H1). H1 was tested using all trainee groups and their matched controls across the data set. A total of 24 matched trainee-control pairs was generated. The trainee and control groups were then compared on concern about gender discrimination, concern about racial discrimination, perceptions of positive inclusion climate, positive attitudes towards workplace diversity, modern sexism, and modern racism. H1 was tested using a series of paired sample *t*-tests. The trainee and control groups were significantly different on concern about gender discrimination (trainee $M = 9.14$, control $M = 7.62$; $t(22) = -2.31$, $p < .05$), perceptions of positive inclusion climate (trainee $M = 3.96$, control $M = 3.07$; $t(21) = -2.82$, $p < .05$), attitudes toward workplace diversity (trainee $M = 4.55$, control $M = 4.23$; $t(21) = -2.67$, $p < .05$), and modern sexism (trainee $M = 1.88$, control $M = 2.47$; $t(22) = 1.99$, $p < .05$). Additionally, trainees had more positive feelings toward diversity (trainee $M = 4.57$, control $M = 4.27$; $t(21) = -2.57$, $p < .05$) and fewer negative feelings toward diversity (trainee $M = 1.48$, control $M = 1.82$; $t(21) = 2.31$, $p < .05$) than the control group. The trainee and control groups were not significantly different on concern about racial discrimination (trainee $M = 9.64$, control $M = 8.78$; $t(22) = -1.58$, $p = .129$) and modern racism (trainee $M = 1.52$, control $M = 1.53$; $t(21) = .024$, $p = .981$). As such, H1 was partially supported.

Hypothesis 2 (H2). H2 proposed that the TIGER Advocates trainees differed from their matched control group on the outcomes of concern about gender discrimination, concern about racial discrimination, perceptions of inclusion climate, attitudes towards workplace diversity, modern sexism, and modern racism. A total of

seven matched trainee-control pairs was generated, and H2 was tested using a series of paired *t*-tests. The trainee and control groups were significantly different on perceptions of positive inclusion climate (trainee $M = 4.29$, control $M = 2.40$; $t(6) = -3.94$, $p < .01$). However, the trainee and control groups were not significantly different on concern about gender discrimination, concern about racial discrimination (trainee $M = 10.00$, control $M = 9.56$; $t(6) = -1.00$, $p = .356$), attitudes toward workplace diversity (trainee $M = 4.46$, control $M = 4.19$; $t(6) = -.731$, $p = .419$), modern sexism (trainee $M = 1.72$, control $M = 2.23$; $t(6) = 1.147$, $p = .30$), and modern racism (trainee $M = 1.14$, control $M = 1.69$; $t(6) = 1.79$, $p = .124$). As such, H2 was partially supported.

Hypothesis 3 (H3). H3 proposed that the Towards Equitable Workloads trainees would differ from their respective control group on the outcomes of concern about gender discrimination, modern sexism, perceptions of inclusion climate, and attitudes toward workplace diversity. However, due to the small sample size, only three matched trainee-control pairs could be generated from the data. Thus, there was insufficient data to test H3.

Hypothesis 4 (H4). H4 proposed that TPR Equitable Evaluation trainees would differ from the control group on concern about gender discrimination, concern about racial discrimination, perceptions of inclusion climate, attitudes towards workplace diversity, modern sexism, and modern racism. A total of seven matched trainee-control pairs could be generated, and H4 was tested using a series of paired *t*-tests. However, the trainee and control groups were not significantly different on any of the outcome variables. As such, H4 was not supported.

Hypothesis 5 (H5). H5 proposed that the Trailblazers program participants would differ from their respective control group on concern about gender discrimination, modern sexism, perceptions of inclusion climate, and attitudes toward workplace diversity. A total of nine matched trainee-control pairs could be generated, and H5 was tested using a series of paired *t*-tests. However, the trainee and control groups were not significantly different on any of the outcome variables. As such, H5 was not supported.

Research Question Analyses Results

Research questions 1 through 6 were exploratory in nature because the impact of COVID-19 on voluntary learning initiatives in higher education has been largely unexplored. The research questions of the present study were primarily concerned with ascertaining whether changes in faculty work hours, childcare responsibilities, and household responsibilities were related to faculty willingness to participate in voluntary learning opportunities during COVID-19, including voluntary trainings, diversity and inclusion initiatives, and continuing education seminars. These research questions were answered primarily through correlational analyses, *t*-tests, and analyses of variance (ANOVAs). However, additional analyses of interest were conducted as well.

General Impact of COVID-19. About 47% of the sample reported that they had significant childcare responsibilities, and of those faculty, 86.4% reported that their childcare responsibilities made it more difficult for them to be productive in their research, and 88.9% reported that their childcare responsibilities made it more difficult to be productive in their teaching. Additionally, about 87% of the sample reported that they had significant household responsibilities. Of those faculty, 72.1% reported that their

household responsibilities made it more difficult to be productive in their research, and 67.2% reported that their household responsibilities made it more difficult to be productive in their teaching. Additionally, COVID-19 appears to have impacted faculty's teaching and research hours. 37.2% of the faculty in the sample reported that their summer teaching hours had increased during COVID-19, while 71.8% reported that their teaching hours during the academic year had increased. A different pattern was seen for research hours. 48.8% reported that their research hours during the summer had decreased, while 62% reported that their research hours during the academic year had decreased.

Research Question 1 (RQ1). RQ1 asked if there was a relationship between significant childcare responsibilities and faculty willingness and ability to participate in voluntary learning opportunities, like voluntary trainings, diversity and inclusion initiatives, and continuing education seminars. Independent samples *t*-tests were conducted with the presence of childcare responsibilities as the grouping variable and faculty responses about their willingness and ability to participate in voluntary learning opportunities as the outcome variables. Compared to people with no childcare responsibilities, those with childcare responsibilities reported lower ability to participate in voluntary learning opportunities (childcare $M = 2.02$, $SD = .926$; no childcare $M = 2.65$, $SD = 1.25$), $t(76) = 2.52$, $p < .05$. However, there were no significant differences between the two groups in terms of the reported effects of COVID-19 on their willingness to participate in voluntary learning opportunities.

Research Question 2 (RQ2). RQ2 asked if there was a relationship between significant household responsibilities and faculty willingness and ability to participate in voluntary learning opportunities, like voluntary trainings, diversity and inclusion initiatives, and continuing education seminars. Independent samples *t*-tests were conducted with the presence of household responsibilities as the grouping variable and faculty responses about their willingness and ability to participate in voluntary learning opportunities as the outcome variables. The presence of household responsibilities had no impact on either faculty willingness or ability to participate in voluntary trainings, diversity and inclusion initiatives, and continuing education seminars.

Research Question 3 (RQ3). RQ3 asked if there was a relationship between changes in faculty's summer research hours during COVID-19 and their ability and willingness to participate in voluntary learning opportunities, including voluntary trainings, diversity and inclusion initiatives, and continuing education seminars. An ANOVA was conducted with changes in summer research hours as the grouping variable and ability to participate in voluntary learning opportunities as the outcome variable. The main effect of changes in summer research hours on ability to attend voluntary learning opportunities was significant, $F(4, 73) = 5.18, p < .01$. Those who reported that their summer research hours decreased significantly ($M = 1.76, SD = .907$) were significantly less able to attend voluntary learning opportunities than those who reported that their summer research hours either remained the same ($M = 2.69, SD = .933$) or increased slightly ($M = 3.37, SD = 1.49$). Similarly, changes in summer research hours significantly impacted willingness to participate in voluntary learning opportunities, $F(4, 73) = 4.71, p$

< .01. Those who reported that their summer research hours decreased significantly ($M = 1.87$, $SD = .948$) were significantly less willing to participate in voluntary learning opportunities during COVID-19 than those who reported that their summer research hours either remained the same ($M = 2.78$, $SD = .952$) or increased slightly ($M = 3.37$, $SD = 1.44$).

Research Question 4 (RQ4). RQ4 asked if there was a relationship between changes in faculty's academic year research hours during COVID-19 and their ability and willingness to participate in voluntary learning opportunities, including voluntary trainings, diversity and inclusion initiatives, and continuing education seminars. An ANOVA was conducted with changes in academic year research hours as the grouping variable and ability to participate in voluntary learning opportunities as the outcome variable. The main effect of changes in academic year research hours on ability to attend voluntary learning opportunities was significant, $F(4, 73) = 7.11$, $p < .001$. Those who reported that their academic year research hours decreased significantly ($M = 1.62$, $SD = .755$) were significantly less able to participate in voluntary learning opportunities during COVID-19 than those who reported that their hours decreased just slightly ($M = 2.57$, $SD = .937$), remained the same ($M = 2.81$, $SD = 1.14$) or increased slightly ($M = 3.27$, $SD = 1.50$). Similarly, changes in academic year research hours significantly impacted willingness to participate in voluntary learning opportunities, $F(4, 73) = 5.98$, $p < .001$. Those who reported that their academic year research hours decreased significantly ($M = 1.86$, $SD = .915$) were significantly less willing to participate in voluntary learning

opportunities during COVID-19 than those who reported that their hours either remained the same ($M = 3.12$, $SD = .950$) or increased slightly ($M = 3.20$, $SD = 1.35$).

Research Question 5 (RQ5). RQ5 asked if there was a relationship between changes in faculty's summer teaching hours during COVID-19 and their ability and willingness to participate in voluntary learning opportunities, including voluntary trainings, diversity and inclusion initiatives, and continuing education seminars. An ANOVA was conducted with changes in summer teaching hours as the grouping variable and ability to participate in voluntary learning opportunities as the outcome variable, but the main effect of changes in summer teaching hours on ability to participate in voluntary learning opportunities was not significant. Another ANOVA was conducted with changes in summer teaching hours as the grouping variable and willingness to participate in voluntary learning opportunities as the outcome variable, and the main effect of changes in summer teaching hours on willingness to participate in voluntary learning opportunities was also insignificant.

Research Question 6 (RQ6). RQ6 asked if there was a relationship between changes in faculty's academic year teaching hours during COVID-19 and their ability and willingness to participate in voluntary learning opportunities, including voluntary trainings, diversity and inclusion initiatives, and continuing education seminars. An ANOVA was conducted with changes in academic year teaching hours as the grouping variable and ability to participate in voluntary learning opportunities as the outcome variable, but the main effect of changes in academic year teaching hours on ability to participate in voluntary learning opportunities was not significant. Another ANOVA was

conducted with changes in academic year teaching hours as the grouping variable and willingness to participate in voluntary learning opportunities as the outcome variable, and the main effect of changes in academic year teaching hours on willingness to participate in voluntary learning opportunities was also insignificant.

DISCUSSION

Discussion of Hypothesis Testing and Results

The purpose of the present study was to explore the impact of four different diversity training programs implemented at a higher education institution, which were designed to help participants recognize bias in the workplace, teach participants methods they could use to confront bias, and help participants prevent those biases from impacting workplace decisions. However, it should be noted that on many of the outcome variables, there may have been floor and ceiling effects present. Across the whole sample, participants appeared to already be relatively high on concern about racial ($M = 9.38$, $SD = 1.49$) and gender discrimination ($M = 8.61$, $SD = 1.8$) and low on modern sexism ($M = 2.05$, $SD = 1.01$) and modern racism ($M = 1.42$, $SD = 0.75$). As such, this may have hindered the ability to detect the effects of training on the outcome variables.

H1 suggested that, compared to their respective control groups, trainees from across all four training programs would be more concerned about gender and racial discrimination, have greater perceptions of inclusion climate, have more positive attitudes toward workplace diversity, and exhibit lower levels of racism and sexism. H1 was partially supported, in that trainees across all programs were significantly more concerned about gender discrimination, endorsed sexist statements at a significantly

lower rate, had greater perceptions of inclusion climate in their departments, and had more positive attitudes toward workplace diversity than the control group. The trainees did not exhibit any difference from the control group in concern about racial discrimination or endorsement of racist statements. Because more of the focal training programs of the study emphasized sexism and gender discrimination issues, perhaps there were no differences in the racism-related outcomes because race issues were simply not emphasized frequently enough in the training materials to impact attitudinal outcomes.

H2 suggested that TIGER Advocates trainees would differ from their matched control group on the outcomes of concern about gender discrimination, concern about racial discrimination, perceptions of inclusion climate, attitudes towards workplace diversity, modern sexism, and modern racism. H2 was partially supported because the trainees had significantly greater perceptions of inclusion climate in their departments and were marginally more concerned about gender discrimination than the control group. While trainees and controls did not differ significantly on any other outcomes, I believe this might have been an issue of power, since there were only seven matched pairs to conduct statistical tests on. Differences between trainees and controls on all of the outcomes were in the expected directions, but there was simply not enough data to support substantial conclusions. This conclusion is further supported because a previous training evaluation of the TIGER Advocates program using a larger sample ($n = 25$ matched pairs) found that trainees were significantly lower than controls in modern sexism and higher than controls in concern about gender discrimination (D'Souza, 2017).

H3 proposed that, compared to their control group, the Towards Equitable Workloads trainees would have greater concern about gender discrimination, lower modern sexism, greater perceptions of inclusion climate, and more positive attitudes toward workplace diversity. However, H3 could not be tested because very few participants in the sample had taken the Towards Equitable Workloads training, and only three matched trainee-control pairs could be generated from the sample. As such, more data would need to be collected to draw reliable conclusions about the effectiveness of this specific training program.

H4 proposed that TPR Equitable Evaluation trainees would differ from their control group on concern about gender discrimination, concern about racial discrimination, perceptions of inclusion climate, attitudes towards workplace diversity, modern sexism, and modern racism. However, H4 was not supported because trainees did not differ from the controls on any of these outcomes. There are several reasons why this might have been the case. First of all, there could be a power issue with this training program as well, since only seven matched pairs could be generated from the sample to test H4. Additionally, the TPR Equitable Evaluation training was administered after the start of COVID-19 in a virtual format. Perhaps the virtual format of the training was simply less effective for inspiring attitudinal change, or participants might have been too distracted by current events to allocate enough attention to the virtual training for it to be effective. This training program should be evaluated again to draw more substantial conclusions, perhaps using a larger sample and an in-person format.

H5 proposed that the Trailblazers program participants would differ from their respective control group on concern about gender discrimination, modern sexism, perceptions of inclusion climate, and attitudes toward workplace diversity. However, H5 was not supported because trainees and controls did not differ significantly on any of these outcomes. There are a few reasons why this could be the case. First of all, the Trailblazers program was more of a leadership initiative than a diversity training program, so the program content may not have emphasized sexism and gender discrimination issues enough to make an impact on these attitudinal outcomes. Additionally, power may be an issue with this program as well because only eight paired samples could be generated from the sample to test H5. Last, 70% of the Trailblazers participants in the sample had begun the program in 2019 or earlier. As such, any attitudinal changes that were a result of the program may have faded by the time the participants took the survey for the present study in 2021.

Contributions to Diversity Training Research and Practice

Although not all of the training programs focal to the present study were successful in impacting all of the outcomes they were expected to, this study still supports the implication that diversity and inclusion trainings and initiatives can successfully affect explicit attitudinal outcomes. Although previous research has found that diversity training has a smaller effect on attitudes than on behaviors and cognitions (e.g., Bezrukova et al., 2016), this is still an important avenue of investigation because the goal of diversity training is often to primarily address attitudinal outcomes, like prejudice, stereotyping, and biases. As such, any additional research evidence showing a

change in attitudinal outcomes due to diversity training is valuable and useful to the literature.

Another important contribution of the present study was the use of propensity score matching methodology. Propensity score analysis is a relatively novel approach in the diversity training literature. A search of the relevant published research literature in June 2021 found only one other example of a study that had used the propensity scoring approach when evaluating a diversity and inclusion intervention – a study by Bowman, Denson, and Park (2016). As shown in the present study, propensity score matching is a unique approach that is very well-suited to evaluating the effectiveness of training programs when random assignment to training is not feasible or possible, as is often the case in organizational settings. The present study demonstrated that when participants are equated across trainee and control conditions using propensity score matching, some notable differences in the expected directions are found. This is in contrast to the more traditional approach of conducting simple independent samples *t* tests comparing all trainees to all controls without regard to underlying (and confounding) differences between the two groups. For example, when I used the propensity score approach to test Hypothesis 1, I found that trainees and controls were significantly different on concern about gender discrimination, modern sexism, perceptions of inclusion climate, and attitudes toward workplace diversity. However, if I had conducted independent samples *t* tests comparing the trainees and controls without using propensity score matching, I would have only found that trainees and controls differed on perceptions of inclusion climate and attitudes toward workplace diversity. The other two significant results I

found using the propensity scoring approach would have otherwise been obscured, theoretically due to the confounds and selection bias innate to the methodological approach of the study (e.g., participants self-selecting into the training and control conditions). Thus, the present study shows that propensity scoring is a useful method for controlling selection bias and other confounds, and this approach can be suitably applied to other organizational settings as well.

This study also gives practitioners in other organizational settings an outline for some diversity management practices that could be used to improve perceptions of organizational inclusion climate and reduce sources of bias at work. Implementing some of these practices could help organizations reap the benefits of a diverse workforce while reducing the personnel costs associated with the turnover, absenteeism, and burnout that can occur with higher levels of workplace diversity.

Contributions to Research on the Impacts of COVID-19

A unique contribution of the present study is that it examined the impact of COVID-19 on participation in voluntary learning opportunities, like diversity training. COVID-19 caused many U.S. workers, including college and university faculty, to start spending more time working remotely (Brynjolfsson et al., 2020), and during the height of the pandemic, many workers were taking on more household and childcare responsibilities than before, due to daycares and schools shutting down. Additionally, because of the changing demands of remote work and household production, many academics were dedicating fewer hours to work tasks, including research and teaching (Myers et al., 2020). The present study partially supported Myers' conclusions, finding

that faculty had to make trade-offs between the time they spent teaching and the time they spent on research during COVID-19. About half (48%) of the faculty in the present study reported having to spend more time on teaching and less time on research during the academic year during the pandemic, and I believe this was due to the demand of having to teach students remotely, as well as in person, due to university policy during COVID-19. Meanwhile, about 22% of the faculty in the study reported decreases in research and increases in teaching during the summer of COVID-19, and a similar percentage (24%) reported no changes in research or teaching hours during the summer of the pandemic. As such, it seems that COVID-19 had less of an impact of faculty's summer semester time allocation than it did on academic year time allocation. The present study sought to determine what impact these work changes caused by COVID-19 had on faculty's ability and willingness to participate in voluntary learning opportunities.

Preliminary research from 2020 found that women faculty and faculty with young dependents were experiencing the most drastic decreases in work hours during COVID-19 (Myers et al., 2020). However, in the sample for the current study, women were no more likely than men to have significant childcare or household responsibilities, nor did men and women report significantly different changes in research or teaching hours during the pandemic. As such, the present study only examined willingness and ability to participate in voluntary learning opportunities by childcare and household responsibilities and changes in working hours, rather than by gender as well. However, it should be acknowledged that women faculty with children may have experienced the impacts of COVID-19 differently than their male colleagues with children. In fact, in the present

study, women faculty with children reported that their childcare responsibilities interfered with their research productivity to a greater extent than male faculty with children did. This finding supports the preliminary findings about the pandemic from Myers and colleagues (2020).

The current study found that changes in faculty's research hours, but not teaching hours, during the summer and the academic year predicted how willing and able they were to participate in voluntary learning opportunities. Faculty who had seen drastic decreases in the time they could dedicate to research during the summer and academic year also reported being less able and willing to participate in voluntary learning opportunities. The present study also found that significant childcare responsibilities, but not household responsibilities, were related to faculty ability, but not willingness, to participate in voluntary learning opportunities. Compared to faculty with no childcare responsibilities, those with childcare responsibilities reported lower ability to participate in voluntary learning opportunities.

In the present study, there was also a connection between substantial decreases in research hours and childcare responsibilities. Having significant childcare responsibilities was negatively correlated with summer research hours during COVID-19 ($r = -.325$). This means that faculty who had childcare responsibilities also saw more drastic decreases in the time they could commit to research over the summer of COVID-19, and this conclusion is supported by a chi-square test, which found that faculty with childcare responsibilities were more likely to report that their summer research hours had decreased significantly compared to faculty with no childcare responsibilities, $X^2(4, n = 79) = 9.632$,

$p = .047$. I believe that this might be due to the lack of typical summer childcare options that were available over the summer of 2020 due to COVID-19 restrictions. As such, faculty juggling childcare responsibilities, along with the other changes in time demands brought on by COVID-19, were most likely to feel that their ability to attend voluntary learning opportunities had decreased due to the pandemic. These findings support preliminary research by Myers and colleagues (2020), who found that faculty with young children were experiencing more drastic decreases in work hours during COVID-19.

In summary, the present study found evidence that the context of COVID-19, a time rapid and drastic change, impacted faculty's willingness and ability to attend voluntary trainings, diversity and inclusion initiatives, and continuing education seminars. As such, this study contributes to the literature on training because it emphasizes the impact that context can have on training participation and motivation.

Limitations and Directions for Future Research

The primary limitation of the present study was the small sample size, especially when attempting to determine the impacts of individual training programs. Although the total sample size for the study was 79, the number of matched pairs generated to test each training program's hypotheses ranged from only three to 24. Clearly, larger sample sizes are needed to draw more definite conclusions about these training programs' effectiveness. Future researchers, perhaps in applied settings, may find it more feasible than the present study to collect large volumes of data to test the effectiveness of these and other diversity training programs using the same propensity scoring method used in this study.

Another limitation of the present study was that the data was not longitudinal. I did not have data comparing the same participants before and after participation in each training program, and this is a well-known limitation of many training evaluation studies (Kalinowski et al., 2013). Additionally, the training and control conditions were not randomly assigned because all the initiatives in the study were voluntary, and this study only used explicit, attitudinal outcome measures. Previous research has found that attitudinal outcomes are more difficult to impact with training than behavioral or cognitive outcomes (Bezrukova et al., 2016; Kalinoski et al., 2013). Given these methodological limitations, future research on diversity training should consider using other research methods (e.g., longitudinal observations and random assignment) and introducing additional outcome measures (e.g., behavioral, cognitive, and implicit attitudinal measures).

Future studies can also expand upon the diversity training methodology research by determining if there are heterogeneous groups of people who may be more responsive or less responsive to different types of training (e.g., behavior-based diversity training versus awareness-based diversity training). Past studies have examined trainee characteristics and their impact on diversity training outcomes, but they have found conflicting results. Bezrukova and colleagues (2016), for example, found that average trainee age, proportion of white trainees, and proportion of female trainees all did not have a significant impact of overall training effect sizes. On the other hand, Kalinoski and colleagues (2013) found that trainee groups with greater proportions of non-white trainees and trainee groups with greater proportions of female trainees were both

associated with larger diversity training effect sizes. Although comparing the differential impact of diversity training across demographically distinct groups was beyond the scope of the present study, future researchers may find it useful to continue to examine the impact of trainee characteristics on diversity training outcomes, while also incorporating the impact of methodological approaches as well.

In the present study, I corrected for non-random training assignment by using propensity score matching to select control groups for each of the training groups based on participants' university college, faculty rank, gender, and race or ethnicity. Future diversity training studies should continue using the propensity scoring approach and should explore using other observed covariates to estimate training propensity scores. In applied settings, these observed covariates might include things like organizational tenure, management level, job title and function, and company department or division. Additionally, one limitation of the present study was that the propensity scoring covariates were observed at the same time as the training outcome variables. Future research should consider measuring the covariates before the outcome variables, as this would help better establish causality. Also, if this method were used, it would allow for pre-training attitudinal variables (e.g., modern sexism and racism, concern about discrimination) to be used as covariates, which would help better establish the impact of training programs.

Last, future research should continue to examine the impact of context on training participation and motivation, as the present study found intriguing results regarding the

impact that COVID-19 had on faculty willingness and ability to participate in voluntary learning opportunities.

Summary of Research Contributions and Practical Implications

In summary, the present study expanded the research literature on diversity training by further examining attitudinal training outcomes, using a novel methodological approach to evaluate the outcomes of diversity training, and studying how the context of COVID-19 and related changes in time demands and responsibilities impacted faculty willingness and ability to participate in optional trainings and initiatives. As for practical implications, this study gives practitioners useful evidence for the effectiveness of diversity training programs, particularly in higher education, and it provides insight into the potential effect of contextual factors (i.e., the COVID-19 pandemic) on training participation and motivation. Additionally, training practitioners in higher education can use the findings of the present study to inform decisions about increasing participation in voluntary training programs. This study found that, obviously, faculty have limited time and many obligations, and participation in voluntary learning opportunities will decline if faculty feel that they have too many other responsibilities that take precedence. As such, training program managers need to emphasize the importance of these trainings and perhaps work with administration to lighten workloads and reduce obligations on days that faculty need to attend important programs. Implementing such policies may help minimize the number of responsibilities faculty are balancing so they will be more likely to attend.

CONCLUSION

Diversity in the workplace can have many benefits for organizations, including increased team creativity, job satisfaction, organizational commitment, longer tenure, and better retention (Li et al., 2017; Mor Barak et al., 2016). However, when organizations fail to properly manage diversity, the detrimental outcomes of diversity, including higher turnover, absenteeism, increased job stress, emotional exhaustion, and depersonalization, can far outweigh the benefits (Mor Barak et al., 2015). Fortunately, many organizations have successfully managed workplace diversity through diversity training and other interventions. Meta-analytic evidence has shown that diversity training can have meaningful and positive effects on cognitive learning, on-the-job behavior, and affective outcomes (Bezrukova et al., 2016; Kalinoski et al., 2013), and the present study supported previous research findings indicating that diversity training can positively impact attitudinal outcomes.

In the past, organizations have simply focused on increasing the representation of women and racial and ethnic minorities in the workforce, but this approach does not consistently improve the workplace. Instead, creating a positive organizational climate for inclusion is of the utmost importance to organizations seeking to improve their workplace through diversity, and implementing effective diversity training is one way to create an inclusive climate and improve organizational outcomes.

REFERENCES

- Arthur, W., Bennett, W., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. *Journal of Applied Psychology, 88*(2), 234 – 245. Doi: 10.1037/0021-9010.88.2.234
- Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research, 46*, 399 – 424. Doi: 10.1080/00273171.2011.568786
- Bartel, C. A., Wrzesniewski, A., & Wiesenfeld, B. M. (2012). Knowing where you stand: Physical isolation, perceived respect, and organizational identification among virtual employees. *Organization Science, 23*(3), 743 – 757. Doi: <http://dx.doi.org/10.1287/orsc.1110.0661>
- Bell, S. T., Villado, A. J., Lukasik, M. A., Belau, L., & Briggs, A. L. (2011). Getting specific about demographic diversity variable and team performance relationships: A meta-analysis. *Journal of Management, 37*, 709–743. Doi: 10.1177/0149206310365001
- Bezrukova, K., Jehn, K. A., & Spell, C. S. (2012). Reviewing diversity training: Where we have been and where we should go. *Academy of Management Learning & Education, 11*(2), 207 – 227. Doi: <http://dx.doi.org/10.5465/amle.2008.0090>
- Bezrukova, K., Spell, C. S., Perry, J. L., & Jehn, K. A. (2016). A meta-analytical integration of over 40 years of research on diversity training evaluation. *Psychological Bulletin, 142*(11), 1227 – 1274. Doi: <http://dx.doi.org/10.1037/bul0000067>

- Bowman, N. A., Denson, N. & Park, J. J. (2016). Racial/cultural awareness workshops and post-college civic engagement: A propensity score matching approach. *American Educational Journal*, 53(6), 1556-1587. Doi: <https://doi.org/10.3102%2F0002831216670510>
- Brouwer, M. A. R., & Boros, S. (2010). The influence of intergroup contact and ethnocultural empathy on employees' attitudes toward diversity. *Cognition, Brain, & Behavior*, 14(3), 243-260.
- Brown, K. N. (2019). *Implicit stereotypes: An explanation for the lack of female leadership in the STEM fields?* (13858584) [Master's thesis, Clemson University]. ProQuest Dissertations and Theses.
- Brynjolfsson, E., Horton, J., Ozimek, A., Rock, D., Sharma, G., & Ye, Y. T. (2020). *COVID-19 and remote work: An early look at U. S. data*. Unpublished manuscript.
- Carnes, M., Devine, P. G., Manwell, L. B., Byars-Winston, A., Fine, E., Ford, C. E., Forscher, P., Isaac, C., Kaatz, A., Magua, W., Palta, M., & Sheridan, J. (2015). The effect of an intervention to break the gender bias habit for faculty at one institution: A cluster randomized, controlled trial. *Academic Medicine*, 90(2), 221 – 230. Doi: 10.1097/ACM.0000000000000552
- Colby, S. L. & Ortman, J. M. (2015). Projections of the size and composition of the U.S. population: 2014 to 2060. United States Census Bureau. <https://www.census.gov/content/dam/Census/library/publications/2015/demo/p25-1143.pdf>

- Cunningham, G. B. (2015). Creating and sustaining workplace cultures supportive of LGBT employees in college athletics. *Journal of Sport Management*, 29, 426 – 442. Doi: <http://dx.doi.org/10.1123/JSM.2014-0135>
- Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. L. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, 48, 1267-1278. Doi: 10.1016/j.jesp.2012.06.003
- D'Souza, K. (2017). [2017 Climate Survey] [Unpublished raw data]. Clemson University.
- Duguid, M. M. & Thomas-Hunt, M. C. (2015). Condoning stereotyping? How awareness of stereotyping prevalence impacts expression of stereotypes. *Journal of Applied Psychology*, 100(2), 343 -359. Doi: <http://dx.doi.org/10.1037/a0037908>
- Dwertmann, D. J. G. & Boehm, S. A. (2016). Status matters: The asymmetric effects of supervisor-subordinate disability incongruence and climate for inclusion. *Academy of Management Journal*, 59(1), 44 – 64. Doi: <http://dx.doi.org/10.5465/amj.2014.0093>
- Giffords, E. D. (2009). An examination of organizational commitment and professional commitment and the relationship to work environment, demographic and organizational factors. *Journal of Social Work*, 9, 386–404. Doi:10.1177/1468017309346232
- Haynes, M. (2020). *Top 10 workplace trends for 2020*. Society for Industrial and Organizational Psychology. <https://www.siop.org/Research-Publications/Items-of-Interest/ArtMID/19366/ArticleID/3361/Top-10-Workplace-Trends-for-2020>

- Herring, C. (2009). Does diversity pay? Race, gender, and the business case for diversity. *American Sociological Review*, 74(2), 208 -224. Doi: <https://doi.org/10.1177/000312240907400203>
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, 42(8), 1 – 28. Doi: <https://www.jstatsoft.org/v42/i08/>
- Kalinoski, Z. T, Steele-Johnson, D., Peyton, E. J., Leas, K. A., Steinke, J., & Bowling, N. A. (2013). A meta-analytic evaluation of diversity training outcomes. *Journal of Organizational Behavior*, 34, 1076 – 1104. Doi: 10.1002/job.1839
- Konrad, A. M., Yang, Y., & Maurer, C. (2016). Antecedents and outcomes of diversity and equality management systems: An integrated institutional agency and strategic human resource management approach. *Human Resource Management*, 55(1), 83 – 107. Doi: 10.1002/hrm.21713
- Kunze, F., Boehm, S., & Bruch, H. (2013). Age, resistance to change, and job performance. *Journal of Managerial Psychology*, 28(7/8), 741 – 760. Doi: <https://dx.doi.org/10.1108/JMP-06-2013-0194>
- Li, C., Lin, C., Tien, Y., & Chen, C. (2017). A multilevel model of team cultural diversity and creativity: The role of climate for inclusion. *The Journal of Creative Behavior*, 51(2), 163 – 179. Doi: 10.1002/jocb.93
- Lindsey, A., King, E., Hebl, M., & Levine, N. (2015). The impact of method, motivation, and empathy on diversity training effectiveness. *Journal of Business Psychology*, 30, 605 – 617. Doi: 10.1007/s10869-014-9384-3

- Luellan, J. K., Shadish, W. R., Clark, M. H. (2005). Propensity scores: An introduction and experimental test. *Evaluation Review*, 29, 530 – 558. Doi: 10.1177/0193841X05275596
- Mor Barak, M. E., Lizano, E. L., Kim, A., Duan, L., Rhee, M., Hsiao, H., & Brimhall, K. C. (2016). The promise of diversity management for climate of inclusion: A state-of-the-art review and meta-analysis. *Human Service Organizations: Management, Leadership, & Governance*, 40(4), 305 – 333. Doi: <http://dx.doi.org/10.1080/23303131.2016.1138915>
- Myers, K. R., Tham, W. Y., Yin, Y., Cohodes, N., Thursby, J. G., Thursby, M. C., Schiffer, P., Walsh, J. T., Lakhani, K. R., & Wang, D. (2020). Unequal effects of the COVID-19 pandemic on scientists. *Nature Human Behavior*, 4, 880 – 883. Doi: <https://doi.org/10.1038/s41562-020-0921-y>
- Nishii, L. H. (2013). The benefits of climate for inclusion for gender-diverse groups. *Academy of Management Journal*, 56(6), 1754 – 1774. Doi: <http://dx.doi.org/10.5465/amj.2009.0823>
- Nishii, L. H. & Mayer, D. M. (2009). Do inclusive leaders help to reduce turnover in diverse groups? The moderating role of leader–member exchange in the diversity to turnover relationship. *Journal of Applied Psychology*, 94(6), 1412 – 1426. Doi: 10.1037/a0017190
- O’Leary, B. J. & Weathington, B. L. (2006). Beyond the business case for diversity in organizations. *Employee Responsibilities & Rights Journal*, 18(4), 283-292. Doi: <https://doi.org/10.1007/s10672-006-9024-9>

- Paolillo, A., Silva, S. A., & Pasini, M. (2016). Promoting safety participation through diversity and inclusion climates. *International Journal of Workplace Health Management, 9*(3), 308 – 327. Doi: 10.1108/IJWHM-01-2015-0002
- Preston, M. & De Graaf, S. (2019). *Benefits of socioeconomic diversity to organizations: How organizations can promote and benefit from socioeconomic diversity*. SIOP White Paper Series.
<https://www.siop.org/Portals/84/docs/White%20Papers/Socioeconomic%20Diversity.pdf>
- Putnam-Walkerly, K. & Russell, E. (2016). *What the heck does 'equity' mean?* Stanford Social Innovation Review.
https://ssir.org/articles/entry/what_the_heck_does_equity_mean
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Rosopa, P., Brown, K., & Xoxakos, P. (2019, April). *(Modern) sexism in STEM: Gender differences in academia*. Poster session presented at the annual meeting of the Society for Industrial and Organizational Psychology, National Harbor, MD.
- Rosopa, P. J., Fynes, J. M., D'Souza, K. B., & Xoxakos, P. (2020). In pursuit of inclusivity: A review of human resource management research and practices involving LGBT employees. In D. L. Stone, J. H. Dulebohn, and K. M. Lukaszewski (Eds.), *Diversity and inclusion in organizations*, pp. 165-192. Charlotte, NC: Information Age.

- Sitzmann, T., Brown, K. G., Casper, W. J., Ely, K. & Zimmerman, R. D. (2008). A review and meta-analysis of the nomological network of trainee reactions. *Journal of Applied Psychology, 93*(2), 280 – 295. Doi: 10.1037/0021-9010.93.2.280
- Society for Human Resource Management. (2020). *Express requests – Top resources for your toughest HR questions: Trending topics*.
<https://www.shrm.org/resourcesandtools/tools-and-samples/exreq/pages/trending-topics.aspx>.
- Swim, J. K., Aikin, K. J., Hall, W. S., & Hunter, B. A. (1995). Sexism and racism: Old-fashioned and modern prejudices. *Journal of Personality and Social Psychology, 68*(2), 199-214.
- Tajfel, H. (1982). *Social identity and intergroup relations*. Cambridge, United Kingdom: Cambridge University Press.
- Vespa, J., Medina, L., & Armstrong, D. M. (2020). *Demographic turning points for the United States: Population projections for 2020 to 2060*. United States Census Bureau.
<https://www.census.gov/content/dam/Census/library/publications/2015/demo/p25-1143.pdf>
- Zhang, Y., Punnett, L., & Gore, R. & CPH-NEW Research Team. (2014). Relationships among employees' working conditions, mental health, and intention to leave in nursing homes. *Journal of Applied Gerontology, 33*, 6–23.
Doi:10.1177/0733464812443085

APPENDICES

Appendix A

Demographics Questionnaire

Now we would like to collect some demographic information. Please answer the following questions to the best of your ability.

How would you describe yourself?

- Female
- Male
- Transgender
- Do not identify as male, female, or transgender

Please enter your four-digit birth year.

Please select your race and/or ethnicity.

- American Indian or Alaska Native
- Asian
- Black or African American
- Hispanic or Latino
- Native Hawaiian or Pacific Islander
- White (non-Hispanic)
- More than one race or ethnicity
- Other (please specify) _____

Which college are you in?

- Agriculture, Forestry and Life Sciences
- Architecture, Arts, and Humanity
- Behavioral, Social and Health Sciences
- Business
- Education
- Engineering, Computing, and Applied Sciences
- Science
- Libraries

What is your academic rank, title, or position?

- Professor
- Associate Professor
- Assistant Professor
- Instructor
- Lecturer
- Other _____

Which department are you primarily affiliated with?

What is your tenure status?

- Tenured
- On tenure track but not tenured
- Not on tenure track

Appendix B

Training Checklist

Which of the following TIGERS ADVANCE trainings or initiatives have you participated in while employed at [the institution]?

- TIGER Advocates
- Towards Equitable Workloads Training
- TPR Committee Equitable Faculty Evaluation Training
- Trailblazers Leadership Program
- None of the above

When did you participate in TIGER Advocates training?

- I have not taken this training.
- Before 2018
- 2018
- 2019
- 2020
- 2021

When did you participate in Towards Equitable Workloads training?

- I have not taken this training.
- Before 2018
- 2018
- 2019
- 2020
- 2021

When did you participate in TPR Committee Equitable Faculty Evaluation training?

- I have not taken this training.
- Before 2018
- 2018
- 2019
- 2020
- 2021

When did you participate in the Trailblazers leadership program?

- I have not taken this training.
- Before 2018
- 2018
- 2019
- 2020
- 2021

Appendix C

COVID-19 Impact Questionnaire

Section 1.

The following questions are answered on a five-point scale. 1 = Decreased Significantly, 2 = Decreased Slightly, 3 = Remained the Same, 4 = Increased Slightly, 5 = Increased Significantly.

1. What impact did COVID-19 have on the amount of time you dedicated to research during the last couple of semesters?
 - a. My research hours during the summer...
 - b. My research hours during the academic year...

2. What impact did COVID-19 have on the amount of time you dedicated to teaching during the last couple of semesters?
 - a. My teaching hours during the summer...
 - b. My teaching hours during the academic year...

Section 2.

The following questions are answered as “Yes” or “No”.

1. Do you have significant childcare responsibilities?
2. Do you have significant household responsibilities?

The following questions are answered on a five-point scale. 1 = Significantly more difficult to be productive, 2 = Slightly more difficult to be productive, 3 = No impact on productivity, 4 = Slightly easier to be productive, 5 = Significantly easier to be productive.

1. (If answered “yes” to childcare responsibilities). During COVID-19, what impact have childcare responsibilities had on your research and teaching productivity?
 - a. Research...
 - b. Teaching...
2. (If answered “yes” to household responsibilities). During COVID-19, what impact have household responsibilities had on your research and teaching productivity?
 - a. Research...
 - b. Teaching...

Section 3.

The following question is answered on a five-point scale. 1 = significantly less willing to participate, 2 = Slightly less willing to participate, 3 = No impact on willingness to participate, 4 = Slightly more willing to participate, 5 = Significantly more willing to participate.

1. What impact has COVID-19 had on your willingness to participate in the following?
 - a. Voluntary trainings
 - b. Diversity and inclusion initiatives
 - c. Continuing education seminars

The following question is answered on a five-point scale. 1 = significantly less able to participate, 2 = Slightly less able to participate, 3 = No impact on ability to participate, 4 = Slightly more able to participate, 5 = Significantly more able to participate.

1. What impact has COVID-19 had on your ability to participate in the following?
 - a. Voluntary trainings
 - b. Diversity and inclusion initiatives
 - c. Continuing education seminars

Appendix D
Modern Sexism
Swim, Aikin, Hall, and Hunter (1995)

The following statements are rated on a seven-point scale, ranging from 1 = Strongly Disagree to 7 = Strongly Agree.

1. Discrimination against women is no longer a problem in the United States.
2. Women often miss out on good jobs due to sexual discrimination.
3. It is rare to see women treated in a sexist manner on television.
4. On average people in our society treat husbands and wives equally.
5. Society has reached the point where women and men have equal opportunities for achievement.
6. It is easy to understand the anger of women's groups in the U. S.*
7. It is easy to understand why women's groups are still concerned about societal limitations of women's opportunities. *
8. Over the past few years, the government and news media have been showing more concern about the treatment of women than is warranted by women's actual experiences.

*Items are reverse scored

Appendix E
Concern about Gender Discrimination
Devine et al. (2012)

The following statements are rated on a 10-point scale, ranging from 1 = Strongly Disagree to 10 = Strongly Agree.

1. I'm not personally concerned about discrimination against women. *
2. People need to stop focusing so much time and energy worrying about gender discrimination. *
3. People make more fuss about discrimination against women than is necessary. *
4. I consider gender discrimination to be a serious social problem.

*Items are reverse scored

Appendix F
Concern about Racial Discrimination
Devine et al. (2012)

The following statements are rated on a 10-point scale, ranging from 1 = Strongly Disagree to 10 = Strongly Agree.

1. I'm not personally concerned about discrimination against Black people. *
2. People need to stop focusing so much time and energy worrying about racial discrimination. *
3. People make more fuss about discrimination against Black people than is necessary. *
4. I consider racial discrimination to be a serious social problem.

*Items are reverse scored

Appendix G
Modern Racism
Swim and colleagues (1995)

The following statements are rated on a seven-point scale, ranging from 1 = Strongly Disagree to 7 = Strongly Agree.

1. Discrimination against Black people is no longer a problem in the United States.
2. It is easy to understand the anger of Black people in America. *
3. Black people are getting too demanding in their push for equal rights.
4. Black people should not push themselves where they are not wanted.
5. Over the past few years, the government and news media have shown more respect to Black people than they deserve.

*Item is reverse scored

Appendix H
Climate for Inclusion
Nishii (2013)

The following items are rated on a scale of 1 = Strongly Disagree to 5 = Strongly Agree.

Please answer the following about your department here at [institution].

1. This department has a fair promotion process.
2. The performance review process is fair in this department.
3. This department invests in the development of all of its employees.
4. Employees in this department receive “equal pay for equal work.”
5. This department provides safe ways for employees to voice their grievances.
6. This department is characterized by a non-threatening environment in which people can reveal their “true” selves.
7. This department values work-life balance.
8. This department commits resources to ensuring that employees are able to resolve conflicts effectively.
9. Employees of this department are valued for who they are as people, not just the jobs that they fill.
10. In this department, people often share and learn about one another as people.
11. This department has a culture in which employees appreciate the differences that people bring to the workplace.
12. In this department, employee input is actively sought.
13. In this department, everyone’s ideas for how to do things are given serious consideration.
14. In this department, employees’ insights are used to rethink and redefine work practices.
15. Top management exercises the belief that problem-solving is improved when input from different roles, ranks, and functions is considered.

Appendix I
Workplace Diversity Scale
De Meuse & Hostager (2001)

The following statements are rated on a five-point scale, ranging from 1 = Disagree to 5 = Agree.

1. I believe diversity is fair.
2. Diversity is stressful to me. *
3. I feel enthusiastic about diversity.
4. Diversity is expensive for organizations. *
5. Diversity leads to harmony in organizations.
6. I feel frustrated with diversity. *
7. I feel hopeful about diversity.
8. I believe that diversity is worthless. *
9. I support diversity efforts within organizations.
10. I withdraw from organizational diversity efforts. *
11. Diversity is rewarding to me.
12. I feel resentful about diversity. *
13. Diversity is an asset to organizations.
14. Diversity leads me to make personal sacrifices. *
15. I participate in organizational diversity efforts.
16. I resist organizational diversity efforts. *
17. I believe that diversity is good.
18. Diversity is unprofitable to organizations. *
19. Diversity is enriching for me.
20. I believe that diversity is unjustified. *

*Items are reverse scored for the aggregate scale, or they may be scored separately as a negative attitude toward diversity scale.

Appendix J

R Markdown Document: Propensity Score Analysis Code

```
---  
title: "Dissertation Data Analysis"  
author: "Katie D'Souza"  
date: "5/28/2021"  
output: html_document  
---
```

Setting up the work space. Because I had so few participants, I cleaned my data prior to import into R. This involved removing any cases where the data for the matching covariates was incomplete, because MatchIt can't process the data like that. I also created any dummy variables necessary right in the Excel file, since MatchIt package can't process data that is non-numeric for the matching covariates. However, all of this could have been done in R as well.

```
```{r}  
setwd("/Users/katiedsouza/Downloads")
library(MatchIt)
library(readxl)
library(writexl)
dissdata <- read_excel("DissData.xlsx")
dissdata <- as.data.frame(dissdata)
```
```

Ok, now for our matchit object, predicting Equitable Workloads participation based on gender, race, faculty rank, and college. It matched 3 out of 4 treated cases.

```
```{r}  
predictEquitWorklds <- matchit(EquitWorkld ~ Gender + Race + Rank +
College, data = dissdata, method = "nearest", caliper = .2, calclosest = T)
summary(predictEquitWorklds)
EquitWorkldsData <- match.data(predictEquitWorklds)
EquitWorkldsData
write_xlsx(EquitWorkldsData, "EquitWorkldsData.xlsx")
```
```

Now we can do independent samples t-tests on EquitWorklds data. Turns out, nothing was significantly impacted by the training, even though mean differences were in the right direction. Not surprising, given the 6-person sample. I conducted the paired sample t-tests that were reported in the paper later with a rearranged data file using SPSS.

```
```{r}  
t.test(EquitWorkldsData$GenderDiscScale ~ EquitWorkldsData$EquitWorkld)
t.test(EquitWorkldsData$ModernSexismScale ~ EquitWorkldsData$EquitWorkld)
t.test(EquitWorkldsData$RaceDiscScale ~ EquitWorkldsData$EquitWorkld)
t.test(EquitWorkldsData$ModRaceScale ~ EquitWorkldsData$EquitWorkld)
t.test(EquitWorkldsData$InclusionScale ~ EquitWorkldsData$EquitWorkld)
t.test(EquitWorkldsData$DiversityScale ~ EquitWorkldsData$EquitWorkld)
t.test(EquitWorkldsData$PositiveDiversity ~ EquitWorkldsData$EquitWorkld)
t.test(EquitWorkldsData$NegativeDiversity ~ EquitWorkldsData$EquitWorkld)
```
```

Now the Tiger Advocates training. Matched 7 out of 16 trainee cases. Significantly affected Perceptions of Inclusion, Modern Racism, Gender Discrimination (approached sig at $p < .10$). I conducted the paired sample t-tests that were reported in the paper later with a rearranged data file using SPSS, which got wildly different results than the independent samples t-tests. The paired sample t-tests found that only perceptions of inclusion climate were different.

```
``{r}
predictAdvocates <- matchit(TigerAdvoc ~ Gender + Race + Rank + College,
data = disssdata, method = "nearest", caliper = .2, calclosest = T)
summary(predictAdvocates)
AdvocatesData <- match.data(predictAdvocates)
write_xlsx(AdvocatesData, "TigerAdvocatesData.xlsx")
t.test(AdvocatesData$GenderDiscScale ~ AdvocatesData$TigerAdvoc)
t.test(AdvocatesData$ModernSexismScale ~ AdvocatesData$TigerAdvoc)
t.test(AdvocatesData$RaceDiscScale ~ AdvocatesData$TigerAdvoc)
t.test(AdvocatesData$ModRaceScale ~ AdvocatesData$TigerAdvoc)
t.test(AdvocatesData$InclusionScale ~ AdvocatesData$TigerAdvoc)
t.test(AdvocatesData$DiversityScale ~ AdvocatesData$TigerAdvoc)
t.test(AdvocatesData$PositiveDiversity ~ AdvocatesData$TigerAdvoc)
t.test(AdvocatesData$NegativeDiversity ~ AdvocatesData$TigerAdvoc)
``
```

Here, I started just using the same "TrainingData" name in my code for ease of use. However, if you wanted to name the matched data set something unique, as I did with the first two examples, then you could do that as well. You can also run analyses directly on the data set in R without exporting the Excel file, if that better suits your needs. Now for the Equitable Faculty Evaluation training. Matched 7 out of 8 trainee cases. Training impacted nothing significantly. Found the same thing in the paired sample t-tests later.

```
``{r}
predictTraining <- matchit(TPREquitEval ~ Gender + Race + Rank + College,
data = disssdata, method = "nearest", caliper = .2, calclosest = T)
summary(predictTraining)
TrainingData <- match.data(predictTraining)
write_xlsx(TrainingData, "TPREquitEvalData.xlsx")
t.test(TrainingData$GenderDiscScale ~ TrainingData$TPREquitEval)
t.test(TrainingData$ModernSexismScale ~ TrainingData$TPREquitEval)
t.test(TrainingData$RaceDiscScale ~ TrainingData$TPREquitEval)
t.test(TrainingData$ModRaceScale ~ TrainingData$TPREquitEval)
t.test(TrainingData$InclusionScale ~ TrainingData$TPREquitEval)
t.test(TrainingData$DiversityScale ~ TrainingData$TPREquitEval)
t.test(TrainingData$PositiveDiversity ~ TrainingData$TPREquitEval)
t.test(TrainingData$NegativeDiversity ~ TrainingData$TPREquitEval)
``
```

Now for the Trailblazers program. Matched 9 out of 10 cases. However, nothing was significantly impacted by the training. Found the same thing with the paired sample t-tests later.

```
``{r}
predictTraining <- matchit(Trailblazers ~ Gender + Race + Rank + College,
data = disssdata, method = "nearest", caliper = .2, calclosest = T)
summary(predictTraining)
```

```

TrainingData <- match.data(predictTraining)
write_xlsx(TrainingData, "TrailblazersData.xlsx")
t.test(TrainingData$GenderDiscScale ~ TrainingData$Trailblazers)
t.test(TrainingData$ModernSexismScale ~ TrainingData$Trailblazers)
t.test(TrainingData$RaceDiscScale ~ TrainingData$Trailblazers)
t.test(TrainingData$ModRaceScale ~ TrainingData$Trailblazers)
t.test(TrainingData$InclusionScale ~ TrainingData$Trailblazers)
t.test(TrainingData$DiversityScale ~ TrainingData$Trailblazers)
t.test(TrainingData$PositiveDiversity ~ TrainingData$Trailblazers)
t.test(TrainingData$NegativeDiversity ~ TrainingData$Trailblazers)
...

```

Now any training compared to no training. Matched 24 out of 30 trainee cases. Training significantly impacted Gender Discrimination, Perceptions of Inclusion, Attitudes toward Diversity, Negative attitudes toward diversity, and Modern Sexism (approached significance at $p < .10$). When I ran the paired sample t-tests on this, I found that trainees across all programs were significantly more concerned about gender discrimination, endorsed sexist statements at a significantly lower rate, had greater perceptions of inclusion climate in their departments, and had more positive attitudes toward workplace diversity than the control group.

```

```{r}
predictTraining <- matchit(TrainingAny ~ Gender + Race + Rank + College,
data = disssdata, method = "nearest", caliper = .2, calclosest = T)
summary(predictTraining)
TrainingData <- match.data(predictTraining)
write_xlsx(TrainingData, "AnyTrainingData.xlsx")
t.test(TrainingData$GenderDiscScale ~ TrainingData$TrainingAny)
t.test(TrainingData$ModernSexismScale ~ TrainingData$TrainingAny)
t.test(TrainingData$RaceDiscScale ~ TrainingData$TrainingAny)
t.test(TrainingData$ModRaceScale ~ TrainingData$TrainingAny)
t.test(TrainingData$InclusionScale ~ TrainingData$TrainingAny)
t.test(TrainingData$DiversityScale ~ TrainingData$TrainingAny)
t.test(TrainingData$PositiveDiversity ~ TrainingData$TrainingAny)
t.test(TrainingData$NegativeDiversity ~ TrainingData$TrainingAny)
...

```

And that's the end of the R code I used for my dissertation analyses. Everything else was in SPSS, on the exported matched datasets generated by Matchit package.