




ASSESSING THE EFFICACY OF A CHARACTER DEVELOPMENT PROGRAM IN NON-TRADITIONAL UNDERGRADUATE STUDENTS

 Samuel L
Beaumont^{1*}

 Carol Pernsteiner²

^{1,2}Grand Canyon University, USA.

¹Email: leebeaumont1992@gmail.com Tel: (434)841-0642

²Email: cpernsteiner@my.gcu.edu Tel: (214)842-0749



(+ Corresponding author)

ABSTRACT

Article History

Received: 19 April 2021

Revised: 25 June 2021

Accepted: 28 July 2021

Published: 11 August 2021

Keywords

Adult Education

Character Education

Defining Issues Test-2

Ethics

Non-traditional students

University.

The purpose of this quantitative ex post facto study was to examine if there is a significant difference in mean moral judgment levels of first year undergraduate students after participating in a character development program at a non-traditional program in a university in the southwestern United States. This study included new non-traditional students matriculating between October 2019 and June 2020. The measurement instrument for the study was the Defining Issues Test-2 (DIT-2). Pre- and post-test DIT-2 archival data was retrieved from the study site for the two assessments. The independent variable was the character education program of study, and the dependent variables were the difference in the mean scores from the DIT-2. Paired *t*-tests were used to determine if a statistical difference existed after character training. The positive mean change in post-conventional thinking scores, P ($M = 2.24$), N2 ($M = 1.70$), were not significantly different than zero (P score $p = .203$, N2 score $p = 0.203$). The positive mean change in scores for the lowest level of moral reasoning, PI, ($M = 0.15$) was not significantly different from zero ($p = 0.912$). The predominant moral schema used by the students, maintaining norms (MN) demonstrated a negative mean change ($M = -1.35$) but it was not significantly different from zero ($p = 0.348$). Therefore, the character program of study at the site did not increase moral reasoning levels. The results are significant, as they help to define what works and what does not in non-traditional student ethical development.

Contribution/Originality: This study is one of very few studies which have investigated the efficacy of a comprehensive character education program in non-traditional undergraduate students. These research findings may provide insights for institutional leaders looking to develop and evaluate high-impact character and ethics programming.

1. INTRODUCTION

The decline in civility, increased corporate scandals, and the realization of shared morality for a functional society have led to a resurgence of interest in character education. Individuals need not only academic and technical skills; they also require an understanding of the ethics and morals and associated acceptable behavior guiding their society. Dating back to Aristotle, educational leaders have long recognized the need for the development of the whole individual (Kristjánsson, 2020). Furthermore, the literature has demonstrated the possibility of developing character through education (Abduloh, 2018). As a result, the society expects modern universities to develop ethical graduates, capable of contributing to society. However, no universal approach to character and ethical development exists in higher education. Designing an effective character program of study is also complicated by diverse life

experiences and student demographics present in the modern undergraduate student body. One cohort of particular concern is the non-traditional student.

Despite the fact that many undergraduate learners are classified as non-traditional students (EAB, 2019), character development in this group has not been studied in-depth. Moral development is possible in adults, but it is a process requiring maintenance and continual training (Schinkel & De Ruyter, 2017). For purposes of this study non-traditional students are defined as students who are older than 25, who are employed full-time, have dependents, have delayed their enrollment in an institution of higher education, are military veterans, or a combination of these categories. The problem statement for this study was premised on the absence of a standardized process for ethics and character instruction across institutions of higher education, to study how ethics education is delivered by numerous different pedagogical and andragogical methods, and how educational interventions vary depending on the institution, academic program, and target population.

A review of the applicable literature revealed a lack of research on character and ethical development in non-traditional students. Responding to this gap, researchers examined the efficacy of a character education program of study over one academic year in non-traditional students attending a university in the southwestern United States. The character education program aimed to instill values, morals, ethics, and virtues to students that would help them to be prepared to contribute to society. All students attending the study site were required to participate in the character program of study, and the individuals represented in the study sample volunteered to take a pre- and post-test assessment that happened between October 2019 and December 2020. The aim of this research study was to determine if the study site's character education program of study resulted in a statistically significant increase in student moral development levels.

2. LITERATURE REVIEW

Society has long used education in formal and informal settings, to pass down morals and ethics to the next generation. According to Aristotle, the pursuit of virtue requires an individual to recognize and determine the best course of action by developing positive character traits. Specifically, he touted the concept of *phronesis*, which is applying practical wisdom to determine the appropriate moral action to a given situation. Furthermore, Aristotle stated that these traits could be obtained through education and interactions with others who possess character (Kristjánsson, 2020). Likewise, Confucius believed in the virtuous individual (Mei-Ching Ng, 2009): Not only did Confucius insist that cultivation of virtue is integral to the development of self, but he also believed that it was of great importance for a good society. For him, virtuous people are necessary for the well-functioning of a society. (p. 5). In *Democracy and Education*, Dewey (1916) pointed out that each stage of life possessed challenges and opportunities for development, and development never truly ends for an individual. An individual cannot learn character development in a vacuum as it requires forging relationships, modeling behavior after those engaged in moral action, and repeating ethical behaviors throughout life. The same concept was echoed by Zurqoni, Retnawati, Arlinwibowo, and Apino (2018) after studying students in senior and vocational high schools. This study concluded role models, interventions through various activities, creating habits, and reinforcement were sound strategies for character development. Also, McGrath (2018) argued a prototype for a character education program includes vital components, such as being school-based, structured, and addressing personal and moral growth. Rockenbach (2020) espoused the need for character development by institutions of higher education in the United States. The number of adults with a college degree had increased significantly since the year 2000. Rockenbach (2020) further discussed how the college experience possessed a longitudinal effect on character development by moving students towards the highest level of moral reasoning, termed post-conventional thinking by Kohlberg (1969). However, based on American society's current state, it appears public discourse and civility have decreased over the same period. A recent survey conducted by Weber Shandwick (2019) found 93% of

Americans believe the lack of civility is a pressing national issue. The survey goes on to state that “the nation’s institutions must play a role” (p. 14) in addressing this situation.

Enrollment by non-traditional students in higher education has increased, representing an understudied demographic in terms of the efficacy of character education programs in this group of students. According to the National Center for Education Statistics, about half of all students enrolled in higher education are above the age of 25 (U.S. Department of Education National Center for Education Statistics, 2017). The Lumina Foundation reported 38% of undergraduate students are above the age of 25, 25% are raising children, and approximately 58% work while attending college (EAB, 2019). This change in student demographics created a gap in the research regarding character education and development and begged the question of how educational interventions influence moral development among non-traditional students. Moral development is possible in adults, but it requires maintenance, since inaction can degrade previously developed good habits (Schinkel & De Ruyter, 2017). Furthermore, the evaluation of character programs requires an investment and commitment from educational leaders to develop graduates prepared to contribute to society (Urban & Trochim, 2017).

Butler, Hammel, and Mascia (2011) studied differences in moral reasoning levels in traditional-aged business undergraduate students, non-traditional business undergraduate students, and MBA students. Findings indicated that moral reasoning scores on the Defining Issues Test-2 (DIT-2) were higher for graduate students as compared to traditional business undergraduate students but found no statistically significant difference between graduate students and the non-traditional undergraduate population. One explanation for this was offered by Chen (2017), who illuminated the reasons for group differences in moral development based on age. This suggests that non-traditional students, for the most part, fail to experience the immersive educational experience, spending most of their time outside the confines of their college or university or at jobs or with families. Non-traditional students possess different motivational reasons for pursuing a degree, and their cumulative life experiences influence how they learn. The sample of this study fits this description as the respondent students spend most of their time outside the confines of the university.

3. THEORETICAL FRAMEWORK

Kohlberg’s Theory of Moral Development and Rest’s four-component model were the theoretical foundations for this study. Kohlberg was influenced by the work of Piaget (1932) and his Cognitive Development Theory, which defined four stages: stage one is the sensorimotor stage, stage two is the preoperational stage, stage three is the concrete operational stage, and stage four is the formal operational stage. Piaget (1932) stated how children understood and interacted with the world that drove their moral makeup. Kohlberg (1969) believed moral development was far more complicated than Piaget’s theory, but he did agree that moral development occurred throughout life. However, Kohlberg believed the stages could occur at various points in an individual’s life and create additional development levels and stages. Kohlberg’s moral development theory comprised three levels: pre-conventional, conventional, and post-conventional, and each level possessed two stages (Kohlberg, 1969). All people passed through moral development in the same order, but not at the same pace. The higher stages of moral development resolved moral issues more satisfactorily than earlier stages (Colby & Kohlberg, 1987).

However, Kohlberg’s theory held a few weaknesses that merits disclosure. The primary issue is that his seminal work was based completely on males and therefore may only apply to the male population. Another factor is the use of theoretical situations and their applicability to real-life experiences. A few writers have raised this issue of the relationship between cognitive disruption and the presence or absence of moral action (Saltzstein, 1983). Despite these drawbacks, Kohlberg contributed significantly to the field of moral development and therefore was deemed appropriate as the theoretical foundation for this study.

Using Kohlberg as a foundation, Rest (1986) designed a four-component model consisting of moral sensitivity, moral judgment, moral motivation, and moral courage. Collectively, these influence moral actions on the part of an

individual. Rest (1986) opted for more permeable stages than Kohlberg, referring to them as schema, applying the terms personal interest, maintaining norms, and post-conventional to represent each schema. This model states that moral development occurs continuously and makes a cumulative effect. The individual's developmental level is explained by how likely the person is to utilize one of the three schemas when making moral decisions (Rest, Narvaez, Thoma, & Bebeau, 1999). Individuals of different groups could also present varying moral judgment levels (Rest, Narvaez, Thoma, & Bebeau, 2000).

3.1 Purpose Statement

The purpose of this quantitative ex post facto study was to examine if there is a significant difference in mean moral judgment levels of first-year non-traditional undergraduate students before and after participating in a character education program in a university in the southwestern United States.

3.2. Research Questions

- RQ₁: To what extent is there a significant difference in the moral judgment levels of first-year non-traditional undergraduate students, as measured by the P and N2 scores from the DIT-2, before and after participating in a character education program in a university in the southwestern United States?
- RQ₂: To what extent is there a significant difference in the PI scores of first-year non-traditional undergraduate students, as measured by the DIT-2, before and after participating in a character education program in a university in the southwestern United States?
- RQ₃: To what extent is there a significant difference in the MN scores of first-year non-traditional undergraduate students, as measured by the DIT-2, before and after participating in a character education program in a university in the southwestern United States?

3.3. Hypotheses

To determine the efficacy of the character education program of study, three hypotheses corresponding to each of the three research questions were developed to determine if the character education program of study resulted in a significant difference in moral judgment levels. The study hypotheses tested were:

1. H₁. There a significant difference in the moral judgment levels of first-year non-traditional undergraduate students, as measured by the P and N2 scores from the DIT-2, before and after participating in a character education program in a university in the southwestern United States.
2. H₂. There is a significant difference in the PI scores of first-year non-traditional undergraduate students, as measured by the DIT-2, before and after participating in a character education program in a university in the southwestern United States.
3. H₃. There is a significant difference in the MN scores of first-year non-traditional undergraduate students, as measured by the DIT-2, before and after participating in a character education program in a university in the southwestern United States.

4. METHODS

4.1 The Character Education Program of Study

The study site supplying the data identifies character development as a core outcome in the institutional mission statement and includes a comprehensive character education program with seven defined character traits. These traits are integrity, leader, servant, steward, communicator, dependable, and agile. The academic calendar is accelerated, with students taking three courses every eight weeks (nine semester credit hours) and completing one academic year in six months, resulting in earning a bachelor's level degree in 28 months and an associate's level degree in 15 months. The study site's character education program is embedded across the curriculum and includes

optional co-curricular activities outside the classroom. The character program possesses seven defined program learning objectives (PLOs) completed throughout the degree program and specific course-level objectives.

Specifically, the character program of study includes a dedicated first term course (three credit hours) focused on character and ethical development and an introductory character boot camp. All other academic courses taken by students during this study (24 credit hours) include instructor-facilitated ethical and character discussions customized to the respective class objectives and the individual academic program. Other pedagogical and andragogical assessments include ethical discussion questions located on the study site's learning management system (LMS) and written assignments that require the learner to incorporate character traits and themes. Minor differences between the academic programs exist due to the need to tailor the character assignments to their respective courses and academic programs. Instructor facilitated engagement varies slightly according to the course and the individual instructor. This program aligns with recommendations from other practitioners. For example, [Bowen, Bessette, and Cham \(2006\)](#) advocated for course syllabi to have an ethics component, the use of ethical case studies in courses, and ethics as "an intentional component of the instruction" (p. 6).

4.2. Instrument

The instrument for this study was the DIT-2, which is offered to researchers by the Center for the Study of Ethical Development at the University of Alabama. This assessment tool measured moral reasoning by using ethical dilemmas to activate moral schemas. Moral reasoning level as calculated by the DIT-2 can be represented in four main scores: The lowest level is represented by the Personal Interest score (PI); the middle and the most common are represented by the Maintaining Norms score (MN); and the highest level of moral reasoning is represented by the post-Conventional score (P, N2). The PI score represents the proportion of selected items that appeal to stage 2 and stage 3. The Stage 2 is characterized by decisions that favor the best interest of the participant and stage 3 focuses on the participant's interest in maintaining relationships with others ([Bebeau & Thoma, 2003](#)). The MN score represents the proportion of selected items that appeal to stage 4 considerations, which focuses on the rule of law, formal structures, and maintaining the status quo ([Bebeau & Thoma, 2003](#)). The P score represents where the participant is on the development scale; or in other words, the proportion of items on the DIT selected by the participant. This corresponds to the post-conventional moral framework when making an ethical decision ([Bebeau & Thoma, 2003](#)). Lastly, the N2 score is quite similar to the P score with some minimal differences. The former is considered to have more validity than the latter due to the prioritization of post-conventional items as well as to the degree these post-conventional items are given higher ratings than personal interest items. In addition, N2 score accounts for the degree in which personal interest items are ranked lower by the participants ([Rest et al., 1999](#)). Each of the four scores can have values between 0–95, where a higher value means higher adherence to what the DIT-2 score represents.

Based on Kohlberg's theory of moral development, the DIT 2 uses the three schemas defined by Rest rather than Kohlberg's six stages ([Rest et al., 1999](#)). The assessment consists of five stories and a series of 12 statements for each story. The participant is asked to rate each statement as "should do," "shouldn't do," or "can't decide," and then rank the top four statements in order of importance. Of the 72 statements, 62 represent a stage of moral development. There are 10 meaningless statements, 5 for stage 2, 17 for stage 3, 19 for stage 4, 16 for stage 5, and 5 for stage 6 ([Van den Enden, Boom, & Brugman, 2018](#)). Scores can range from 0 to 95 with moral reasoning levels of <27 considered low, 28–41 considered in the middle, and >41 as high ([Bebeau & Thoma, 2003](#)).

The DIT-2 has been used in over 400 published articles ([Rest et al., 1999](#)) and demonstrates that educational level is responsible for 30–50% of the difference seen in the scores. Seven different criteria assisted in the validation of the instrument. These include age and education level, longitudinal gains, cognitive capacity, moral education interventions, pro-social behaviors, and political leanings. It is reliable with a Cronbach alpha in the upper .7 and

lower .80s (Bebeau & Thoma, 2003); it is valid for both males and females (Rest et al., 1999) and shows no overlap with verbal ability and general intelligence (Thoma, Narvaez, Rest, & Derryberry, 1999).

4.3. Population and Sample

All students at the research site were required to partake in the character education program, but participation in this study was voluntary. We invited all new incoming undergraduate students, who had joined between October 2019 and June 2020, to participate in the pre- and post-test survey of this study. It may be noted that the last pre-test survey was conducted in December 2020, six months after the last cohort (June 2020) that was invited for pre-test survey. No other criteria were used to select the sample, and invitations were sent to 525 students to identify the sample. For the purposes of this study, non-traditional students are defined as those who were married, military veterans, over the age of 24, or a combination of these factors.

4.4. Pre-Test Data Collection

All students were required to take the character education program upon matriculation and so they were also invited to do the pre-test survey upon matriculation. The character education program analyzed in this study had 27 semester credit hours to be completed in a duration of six months. New cohorts of students started every two month and therefore the time of pre-test surveys for each cohort varied from one another. For instance, for those students who matriculated in October 2019, their pre-test survey was also conducted in October 2019; while those who matriculated in December 2019, they also took their pre-test survey in the same month.

The study site's IRB process authorized the institution to collect research data using DIT-2, and students received the instrument through an online survey using SurveyMonkey®. The results were received by the study site and sent to the Center for the Study of Ethical Development at the University of Alabama. Scores were returned to the researchers and data was transcribed into a master Excel® data sheet. The information contained a unique student identification number, the DIT-2 scores (PI, MN, P, N2), and supplemental data retrieved from the study site's student information system (SIS) including the student's age, gender, program of study, and military veteran status.

4.5. Educational Intervention

Upon matriculation, students participated in a six-hour character education program offered by the Director of Institutional Culture and Ethics. Students were provided an outline of the institution's character education program, the importance of investing in ones' character, and several group projects in addition to interactive discussion. The program also contained an experiential learning component. Students were provided a \$10 bill and instructed to give it to someone they encountered in their daily lives who could benefit from the gift. Students then wrote and reflected on their experience.

In the first academic term students were enrolled in a designated three credit hour character and ethics course in addition to two other courses. The objective of the character and ethics course was to further develop a student's understanding of the character education program and how an individual's character and ethics play a role in judgment and decision-making. Specifically, topics included discussing types of ethical situations faced in the global marketplace, cultural differences, the ethical decision-making framework, and the seven specific character traits of the institution. Students also completed an additional 24 credit hours over three academic terms as part of the core character education program. While individual courses varied according to degree program, the character program of study was embedded throughout the academic curriculum and delivered through a combination of discussion questions answered on the school's LMS, classroom-based discussions, and written assignments. Optional co-curricular activities were also offered in the form of lunch meetings with discussions on how to live better; in Bible

study groups; to serve in community service projects; and submitting papers on how to enhance the character program.

4.6. Post-Test Data Collection

Upon the completion of three academic terms (27 semester hours), the volunteers were assessed again using the DIT-2. Each of the cohort had different timings when to complete the post-test survey. For instance, for those who were matriculated in October 2019, their post-test survey was conducted in April 2020. Similarly, students who matriculated in December 2019, accomplished their post-test survey in June 2020. Similar to the pre-test survey, the post-test survey was sent via SurveyMonkey. Results were collected and sent for analyzing their scores. The DIT-2 scores from the post-test were then matched with the pre-test scores in Excel. The student cumulative GPA after completing the 27 semester credit hours was also retrieved from the school's SIS. After the pre and post-test, the information was paired and all personal identifiers were removed, and the data was prepared for statistical analysis in SPSS.

5. RESULTS

The sample consisted of 84 first-year non-traditional undergraduate students who matriculated between October 2019 and June 2020. The response rate for students completing both surveys was 20% among 84 non-traditional students. Paired *t*-tests were used to analyze whether there was a mean significant difference in mean moral judgment scores. The paired *t*-test is the preferred method to ascertain a statistically significant mean difference between paired observations (Xu et al., 2017). The paired *t*-test was appropriate as this study contained one group of individuals measured at two points in time after a treatment effect. Furthermore, the dependent variable was continuous and there was one independent variable consisting of matched pairs (Salkind, 2014). A paired samples *t*-test assumed that the raw numerical data were collected from matched pairs of individuals, each of whom was independent of each other, and that the dependent variable measured at the continuous (interval/ratio) level was normally distributed with no outliers (Field, 2013). These assumptions were satisfied. Table 1 presents the skewness statistics close to zero, indicating that the distributions of the four dependent variables were approximately symmetrical. The Shapiro-Wilk tests indicated no significant deviation from normality ($p > .05$), and there were no outliers because the minimum and maximum Z scores were well within the expected normal limits of ± 3 (Field, 2013).

Table-1. Tests for Skewness and Outliers in Dependent Variables (N = 84).

Dependent variable	Skewness	Minimum Z score	Maximum Z score	Shapiro-Wilk <i>p</i>
P (Post-test minus Pre-test)	0.06	-2.26	2.85	0.730
N2 (Post-test minus Pre-test)	-0.22	-2.79	2.32	0.786
PI (Post-test minus Pre-test)	0.03	-2.52	2.34	0.859
MN (Post-test minus Pre-test)	-0.01	-2.18	2.08	0.250

5.1. Descriptive Findings

This study's target population was non-traditional first-year undergraduate students enrolled in a non-traditional academic program at a university in the southwestern United States. All students at the study site participated in the character development program of study. Table 2 summarizes the demographic and academic characteristics of the sample. The majority of the 84 students ($n = 58, 69.0\%$) were male. They ranged widely in age from 21 to 57 years old with a skewed distribution (Mdn = 28.0 years) in which the majority ($n = 47, 56.0\%$) were 21 to 30 years old. Over half of the students ($n = 47, 56.0\%$) were not veterans. Their academic programs in order of student frequency were Aviation, Information Technology, Nursing, and Business. Their term 3 cumulative GPA

grades ranged widely from 2.17 to 4.00 ($M = 3.57$, $SD = 0.47$) with the majority of students ($n = 47$, 56.0%) achieving grades from 3.1 to 3.9.

Table-2. Demographic and Academic Characteristics of the Sample ($N = 84$).

Characteristic	Category	<i>n</i>	Percent
Gender	Male	58	69.0
	Female	26	31.0
Age (Years)	21 to 30	47	56.0
	31 to 40	26	31.0
	41 to 50	10	11.9
	> 50	1	1.2
Veteran	No	47	56.0
	Yes	37	44.0
Academic Program	Aviation	34	40.5
	Information Technology	27	32.1
	Nursing	12	14.3
	Business	11	13.1
Cumulative GPA	2.1 to 3.0	13	15.5
	3.1 to 3.9	47	56.0
	4.0	24	28.6

Table 3 presents the parametric descriptive statistics (mean and standard deviation) for the four normally distributed dependent variables (PI, MN, P, N2) used to address the research questions, including the pre-test scores, the post-test scores, and the changes in the scores (post-test minus pre-test) for the 84 students who participated in the non-traditional character development program. The mean post-test scores for P, N2, and PI were greater than the corresponding mean pre-test scores; therefore, the mean changes were positive. The mean post-test scores for MN were less than the corresponding mean pre-test scores; therefore, the mean change was negative. The most significant mean change ($M = 2.24$) was for P and the smallest mean change was for PI ($M = 0.15$).

Table-3. Descriptive Statistics for Non-Traditional Student Dependent Variables ($N = 84$).

Dependent variable	Pre-test		Post-test		Post-test minus Pre-test	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
P	24.30	14.47	26.55	13.37	2.24	16.02
N2	21.17	13.82	22.87	12.70	1.70	11.56
PI	27.93	11.91	28.08	11.86	0.15	12.75
MN	39.19	11.50	37.84	12.42	-1.35	13.11

Note: ** Post-conventional scores = P and N2; Personal interest scores = PI; Maintaining norms scores = MN.

Table 4 presents a matrix of Pearson's r correlation coefficients between the post-test minus pre-test changes in the P, N2, PI, and MN scores, as well as the changes in two other variables (Human Liberalism and Religious Orthodoxy) that were in the university database but were not included in the research questions. The correlation matrix indicated that the changes in P, N2, PI, and MN between the pre-test and the post-test were significantly correlated with each other (Pearson's r (84) = $-.541$ to $+.808$, $p < .01$). However, P, N2, PI, and MN were not significantly correlated with the changes in Human Liberalism or Religious Orthodoxy.

Five statistically significant ($p < .01$) negative correlation coefficients were discovered. When the scores for P, N2, and MN increased, then the scores for PI decreased. When the MN scores increased, then the scores for P and N2 decreased. One statistically significant ($p < .01$) positive correlation coefficient was found, when the N2 scores increased, then the P scores increased. The nine correlation coefficients between the changes in the PI, P, N2, and MN scores and the changes in the RO and HL scores were too small to be statistically significant.

Table-4. Correlation Matrix (N = 84).

	P	N2	PI	MN	Human Liberalism	Religious Orthodoxy
P	1					
N2	+0.808**	1				
PI	-0.541**	-0.432**	1			
MN	-0.476**	-0.381**	-0.302**	1		
Human Liberalism	-0.134	-0.105	+0.094	-0.027	1	
Religious Orthodoxy	+0.213	+0.265	-0.245	+0.003	-0.079	1

Note: ** Statistically significant correlation ($p < .01$).

Table 5 presents the results of the paired t -tests. Since p values were larger than .05 for all four of the tests, it was concluded that no significant differences at the .05 level were found between the pre-test and post-test scores. The mean differences (post-test minus pre-test) reported in Table 5 were not significantly different from the differences from zero that could be expected by random chance. The corresponding effect sizes were small (Cohen's $d < 0.2$).

Table-5. Results of Paired Samples t -tests.

Dependent variable	t	df	p	Cohen's d
P	1.28	83	0.203	0.16
N2	1.35	83	0.181	0.13
PI	0.11	83	0.912	0.01
MN	0.94	83	0.348	-0.11

Note: ** Post-conventional scores = P and N2; Personal interest scores = PI; Maintaining norms scores = MN.

Statistical tests for validity included correlation analysis and factor analysis using multiple measures of the same variables. The internal consistency reliability of one variable was estimated using Cronbach's alpha. Test-retest reliability was measured using correlation coefficients between two or more repeated measures. The internal reliability for the DIT-2 was calculated from the story scores for each index, which were supplied in the "raw" SPSS file from the Center for the Study of Ethical Development (Bebeau & Thoma, 2003). However, only demographic data, a case processing summary, and the composite scores for each index were included in the data collection plan and retrieved by the investigator. Therefore, a Cronbach's alpha could not be calculated due to the lack of the story scores for each index (PI, MN, P, N2) for each testing period. When the differences between the means were divided by the standard deviation, the effect size was found to be small (Cohen's $d = 0.2$). Utilizing this effect size, the post-hoc G*Power analysis resulted in a power of .44. Therefore, the study was underpowered, and a minimum sample of 199 students would be required to detect an effect of this size.

1. Statistical evidence using a paired t -test with no violation of assumptions was provided to address RQ₁: *To what extent is there a significant difference in the moral judgment levels of first-year non-traditional undergraduate students, as measured by the P and N2 scores from the DIT-2, before and after participating in a character education program in a university in the southwestern United States?* The positive mean change in scores for P ($M = 2.24$) was not significantly different from zero ($t(83) = 1.28, p = 0.203$). The positive mean change in scores for N2 ($M = 1.70$) was also not significantly different from zero ($t(83) = 1.35, p = 0.203$).
2. Statistical evidence using a paired t -test with no violation of assumptions was provided to address RQ₂: *To what extent is there a significant difference in the PI scores of first-year non-traditional undergraduate students, as measured by the DIT-2, before and after participating in a character education program in a university in the southwestern United States?* The positive mean change in scores for PI ($M = 0.15$) was not significantly different from zero ($t(83) = 0.11, p = .912$).
3. Statistical evidence using a paired t -test with no violation of assumptions was provided to address RQ₃: *To what extent is there a significant difference in the MN scores of first-year non-traditional undergraduate students, as measured by the DIT-2, before and after participating in a character education program in a university in the*

southwestern United States? The negative mean change in the scores for MN ($M = -1.35$) was not significantly different from zero ($t(83) = 0.94, p = .348$). The effect sizes were small (Cohen's $d = -0.11$ to 0.16) for all three research questions. Consequently, it can be concluded that the first academic year of the character education program of study had little effect on character development as measured using the DIT-2.

6. DISCUSSION

A prosperous society requires its members to possess the capability of ethical decision-making for the common good. Throughout history, education has been used as a medium from which society's virtues and values have been passed on to future generations (Langer, Hall, & McMartin, 2010). Higher education leaders realized specialized curriculums and pursued economic advancement as a primary outcome of a college education. This left a character gap in graduates (Yanikoski, 2004). Furthermore, higher education institutions had become far more diverse, including the emergence of a sizeable non-traditional student population (EAB, 2019). Adult students have many competing priorities such as family and careers. These students are often commuters and do not have the same college experience as their traditional-aged counterparts. Non-traditional students enter college with a wealth of life experience resulting in a more defined worldview. According to Thomas and Dunphy (2017), traditional education influences do not produce the same results in the non-traditional student population. The immersion necessary to develop an individual morally and ethically and move them from the personal interest schema to the post-conventional schema in this population is an area of ongoing concern.

The literature demonstrated that as educational level increases, so does post-conventional reasoning, which is represented by the P and N2 scores. Bebeau and Thoma (2003) found that first-year college students present with an average P score of 32.32, and an N2 score of 31.05, while sophomores have slightly elevated levels with P scores of 32.62 and N2 scores of 31.24 and O'Flaherty and Gleeson (2017) discovered similar findings. In a longitudinal study of students in Ireland, P scores increased from 25.37 to 33.16, and N2 scores increased 10.9 points throughout the four-year academic program. In this study, P scores increased from 24.3 to 26.55 ($M=2.24$), and N2 scores increased from 21.17 to 22.87 ($M=1.7$). However, our p values were larger than .05 for the P score ($p=0.203$) and the N2 score ($p=.181$). Therefore, no significant difference occurred at the .05 level, and we accepted the null hypothesis. This finding is not unique as not all ethical and character training interventions resulted in increased post-conventional thinking. Pharmacy students enrolled in a four-module ethics series did not exhibit increased scores on the DIT-2 after training (DeMella et al., 2017). Likewise, a business ethics course in an Indonesian university failed to increase P scores on the DIT (Rafinda, Gál, & Purwaningtyas, 2019).

According to Bebeau and Thoma (2003), the lowest form of moral reasoning represented by PI scores would drop from 28.53 as a freshman to 24.8 as a senior. However, the mean scores for the PI level for this study, as determined by the paired *t*-test, increased slightly from 24.3 to 26.55 ($M=0.15$). There was no statistically significant difference before and after character development training ($p=0.912$) at the .05 level, resulting in the investigator accepting the null hypotheses. This result differs from other studies. Sorensen, Miller, and Cabe (2017) discovered that after participating in an ethics course, the test group's interest scores dropped more than the control group. This finding indicated that students were moving away from the lowest moral reasoning level to the higher levels. Similar results were discovered by O'Flaherty and Gleeson (2017) when PI scores dropped from 37.56 to 31.88 throughout a four-year academic program. Williams (2010) discovered that female interest scores dropped while male scores remained the same after training in the Military Police Corps.

The PI scores for this study were lower than the PI scores reported by Williams (2010) sample, indicating students at the study site utilized this schema at a lower rate. Furthermore, the pre-test PI score for the sample, 24.3, was close to the vocational student's average PI score, 24.87 reported by Bebeau and Thoma (2003). The study site's post-test PI score of 26.55 was slightly higher than the PI scores of junior college students 26.27 as reported by Bebeau and Thoma (2003). The higher PI score may be explained by the percentage of Associates of Applied

Sciences of Airframe and Powerplant (A.A.S.) students in the sample (40.5%), and the fact males comprised the majority (69%) of the sample.

Items on the DIT-2, in which a participant ranks the highest, indicate the preferred schema, allowing for the predominant schema's determination (Rest et al., 2000). The predominant moral schema for the sample in both the pre- and post-tests was maintaining norms. Participants scored 39.19 on the pre-test, and their MN scores dropped to 37.87 on the post-test ($M = -1.35$), but this was not statistically significant ($p = 0.348$). Therefore, we accepted the null hypothesis that no significant difference occurred in the predominant moral schema. Furthermore, the MN (39.19 and 37.87) score exceeded the second most used schema PI (27.93 and 28.08) during both testing periods. The prevalence of maintaining norms as the predominant schema reflects the results of other researchers. O'Flaherty and Gleeson (2017) found that participants demonstrated no change in the MN schema throughout a four-year academic program. However, this study only analyzed the first academic year. Therefore, moral development in the sample retrieved for this study may occur over the remainder of the participant's academic career and their continued participation in the study site's character education program of study. In addition to this, (Maneesong, Wilhelm, Gunawong, & Rachusanti, 2017) discovered that student's predominant moral schema in Thailand was at the MN level, concluding the Thai educational system did not develop post-conventional thinking in students.

Moral reasoning scores also vary by profession and academic program of study. For example, journalism students possessed a P score of 31.18, business students 37.4, and accounting students 42.8 (Auger & Gee, 2016). Many participants in this study were majoring in the A.A.S. Aviation program representing 40.5% of the study sample. This degree differs from the others at the study site as it is more technical, hands-on, and resembles more of a vocational degree than the more traditionally academic-oriented programs at the site. The average P score of a vocational student is 32.19, and the N2 score is 28.7. These are lower than the typical college freshmen who present with a P score of 32.32 and a N2 score of 31.05, and college sophomores who present with P scores of 32.62 and a N2 scores of 31.24 (Bebeau & Thoma, 2003).

This study's problem space remains, requiring additional research on character education efficacy in non-traditional higher education students. This study did not find evidence of moral reasoning development after participating in the character education program. Based on these results, the null hypotheses for each of the three research questions were accepted. A post-hoc G*Power analysis revealed the study was underpowered (.44), and the sample of 84 participants failed to overcome possible Type II errors. Therefore, additional research in this area, accounting for the limitations and recommendations outlined in this study, and studies with larger samples, will advance knowledge of character education in non-traditional post-secondary populations.

7. LIMITATIONS

Limitations of Sampling Strategy

- The study site providing the archival data utilized a convenience sample and only 20% of students who matriculated throughout the study fully completed both assessments. While every effort was made to ensure a representative sample of the school population, the sample may not represent the study site's student body as a whole.

Limitations of sample

- The study was underpowered (0.44), and a sample of 199 would be required to detect a small effect size.
- The small sample did not allow for statistical group comparisons between academic programs, gender, veteran status, and age group.

Limitations of a single study site and impact of COVID-19

- This study's participants were primarily non-traditional students enrolled in a non-traditional academic setting and exposed to a unique character education program of study. The ability to generalize these results to a wider audience was limited.
- The non-traditional academic setting is an accelerated program utilizing eight-week terms, and all students are required to be full-time, taking three courses a term. Therefore, the students are exposed to an educational experience not typically found in other institutions of higher education.
- The unanticipated confounding variable, for which there is no literature available, is the role the global pandemic played in students' development during this study. Educational protocols and methods were changed to reflect the need for social distancing and maintain compliance with local and state restrictions. Most of the data were collected, especially the post-test assessment, during the COVID-19 pandemic. While students were active in pursuing their degrees, several academic terms were completed virtually for all programs. After several months, the aviation technician students and nursing students returned to campus in carefully socially distanced classes for labs and other hands-on experiences. The disruption in the educational process and the strain of everyday life on the non-traditional students may impact the findings of this study, and may not be fit to generalize the findings to a wider audience, nor it can accurately reflect the efficacy of the character education program of study.

8. RECOMMENDATIONS FOR FUTURE RESEARCH

- It is recommended to repeat the study at the same institution with a larger sample to generate sufficient impact and enable group comparisons.
- The DIT-2 does not measure the increase, decrease, or presence of moral actions by participants. The DIT-2 assessment instrument should be compared against a behavioral measurement to determine if there is a relationship between moral development levels and actual moral action. A character education program of study may increase moral judgment levels but not the actual moral acts.
- The character education program of study should also be assessed qualitatively or with a mixed-method research design. Students may have experienced personal growth and reflections that are hard to measure through quantitative methods alone.
- The study site, and other institutions, should also consider investing in a program to train their instructors to ensure the faculty are skilled in delivering ethical instruction and assessing the training program's efficacy.

9. CONCLUSION

While the results did not reject the null hypotheses, the findings have theoretical and practical implications. Despite the limitation of researching moral reasoning growth at only one institution and during a global pandemic, the results of this study provide a springboard for additional future research to determine what works and what does not work in character and ethics instruction in non-traditional students and minority-serving institutions. The character program at the study site is also unique in terms of content and delivery. Unless it is closely replicated, this longitudinal study's results may not be reproducible on other student populations. Therefore, generalizing the results of this study to other schools and student populations may not be applicable. Also, the role demographics play in character development in this population is not fully understood. Some student populations may require different educational interventions and a one-size-fits-all approach may not produce the ethical and moral growth desired by higher education institutions with a diverse student body and a large population of non-traditional students. Therefore, an ongoing need exists to examine different populations, different academic settings, and different educational interventions.

Funding: This study received no specific financial support.

Competing Interests: The authors declare that they have no competing interests.

Acknowledgement: Both authors contributed equally to the conception and design of the study.

REFERENCES

- Abduloh. (2018). *Character education based on six value system*. Paper presented at the SHS Web of Conferences, 00055.
- Auger, G. A., & Gee, C. (2016). Developing moral maturity: An evaluation of the media ethics course using the DIT-2. *Journalism and Mass Communication Educator*, 71(2), 146–162. Available at: <https://doi.org/10.1177/1077695815584460>.
- Bebeau, M. J., & Thoma, S. J. (2003). *Guide for DIT-2*. Minneapolis: Center for the Study of Ethical Development. Minneapolis, MN: University of Minnesota.
- Bowen, C., Bessette, H., & Cham, T. C. (2006). Including ethics in the study of educational leadership. *Journal of College and Character*, 7(7), 1–8.
- Butler, M. N., Hammel, L. R., & Mascia, S. M. (2011). *Utilizing moral reasoning assessment to inform character education in college level business programs*. Paper presented at the Proceedings of the Northeast Business & Economics Association.
- Chen, J. C. (2017). Nontraditional adult learners: The neglected diversity in postsecondary education. *Sage Journals*, 7(1), 1–12. Available at: <https://doi.org/10.1177/2158244017697161>.
- Colby, A., & Kohlberg, L. (1987). *The measurement of moral judgment: Theoretical foundations and research validation* (Vol. 1). Cambridge: Cambridge University Press.
- DeMella, C., Donohoe, K., Morgan, L., Phipps, L., Drisaldi, A., & Forder, M. (2017). Implementation of an online ethics course into a pharmacy curriculum *Currents in Pharmacy Teaching and Learning*, 9(5), 835–840. Available at: <https://doi.org/10.1016/j.cptl.2017.05.001>.
- Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. New York: MacMillan.
- EAB. (2019). Adult learners: Who they are & what they want from college. Retrieved from: <https://www.eab.com/daily-briefing/2019/01/08/adult-learners-who-they-are-and-what-they-want>.
- Field, A. P. (2013). *Discovering statistics using SPSS* (4th ed.). Thousand Oaks, CA: Sage.
- Kohlberg, L. (1969). Stage and sequence: The cognitive developmental approach to socialization. In D. Goslin (Ed.), *Handbook of Socialization Theory and Research* (pp. 347–480). Chicago, IL: Rand McNally.
- Kristjánsson, K. (2020). An introduction to the special issue on wisdom and moral education. *Journal of Moral Education*, 49(1), 1–8. Available at: <https://doi.org/10.1080/03057240.2019.1705041>.
- Langer, R., Hall, M. E. L., & McMartin, J. (2010). Human flourishing: The context for character development in Christian higher education. *Christian Higher Education*, 9(4), 336–360. Available at: <https://doi.org/doi:10.1080/15363759.2010.492738>.
- Maneesong, O., Wilhelm, W., Gunawong, P., & Rachusanti, S. (2017). Survey of ethical reasoning at Chiang Mai University faculty of business administration: Support for international ethics education standards. *Journal of Institutional Research in South East Asia*, 15(3), 19–31.
- McGrath, R. E. (2018). What is character education? Development of a prototype. *Journal of Character Education*, 14(2), 23–36.
- Mei-Ching Ng, R. (2009). College and character: What did Confucius teach us about the importance of integrating ethics, character, learning, and education? *Journal of College and Character*, 10(4), 1–7.
- O’Flaherty, J., & Gleeson, J. (2017). Irish student teacher’s levels of moral reasoning: Context, comparisons, and contributing influences. *Teachers and Teaching Theory and Practice*, 23(1), 59–77. Available at: <https://doi.org/10.1080/13540602.2016.1203777>.
- Piaget, J. (1932). *The moral judgement of the child*. London: Routledge and K. Paul.
- Rafinda, A., Gál, T., & Purwaningtyas, P. (2019). Business ethics course on student moral reasoning. *Oradea Journal of Business and Economics*, 4(special), 60–68.
- Rest, J. R. (1986). *Moral development: Advances in research and theory*. New York: Praeger.

- Rest, J. R., Narvaez, D., Thoma, S. J., & Bebeau, M. J. (1999). DIT2: Devising and testing a revised instrument of moral judgment. *Journal of Educational Psychology*, 91(4), 644–659.
- Rest, J. R., Narvaez, D., Thoma, S. J., & Bebeau, M. J. (2000). A neo-Kohlbergian approach to morality research. *Journal of Moral Education*, 29, 381–395. Available at: <https://doi.org/10.1080/03057240020015001>.
- Rockenbach, A. N. (2020). Character education for the public good: The evolution of character capacities in and beyond college. *Journal of College and Character*, 21(1), 6–13. Available at: <https://doi.org/10.1080/2194587X.2019.1696834>.
- Salkind, N. J. (2014). *Statistics for people who (think they) hate statistics* (5th ed.). Los Angeles: SAGE Publications, Inc.
- Saltzstein, H. D. (1983). Critical issues in Kohlberg's theory of moral reasoning. *Monographs of the Society for Research in Child Development*, 48(1/2), 108–119. Available at: <https://doi.org/10.1111/1540-5834.ep11861811>.
- Schinkel, A., & De Ruyter, D. J. (2017). Individual moral development and moral progress. *Ethical Theory and Moral Practice*, 20(1), 121–127.
- Sorensen, D. P., Miller, S. E., & Cabe, K. L. (2017). Developing and measuring the impact of an accounting ethics course that is based on the moral philosophy of Adam Smith. *Journal of Business Ethics*, 1, 175–192. Available at: <https://doi.org/10.1007/s10551-015-2656-7>.
- Thoma, S. J., Narvaez, D., Rest, J., & Derryberry, P. (1999). Does moral judgment development reduce to political attitudes or verbal ability? Evidence using the Defining Issues Test. *Educational Psychology Review*, 11(4), 325–341.
- Thomas, J., & Dunphy, S. (2017). Factors affecting moral judgment in business students. *Journal of the Indiana Academy of the Social Sciences*, 17(1), 130–153.
- U.S. Department of Education National Center for Education Statistics. (2017). *Table 15: Projections of education statistics to 2025*. Washington, DC: National Center for Education Statistics.
- Urban, J. B., & Trochim, W. M. (2017). Advancing evaluation of character building programs. *Journal of Character Education*, 13(2), 13–27.
- Van den Enden, T., Boom, J., & Brugman, D. (2018). Stages of moral judgment development: Applying item response theory to Defining Issues Test data. *Journal of Moral Education*, 48(4), 1–16.
- Weber Shandwick. (2019). Civility in America 2019: Solutions for tomorrow. Retrieved from: <https://www.webershandwick.com/wp-content/uploads/2019/06/CivilityInAmerica2019SolutionsforTomorrow.pdf>.
- Williams, K. R. (2010). An assessment of moral and character education in initial entry training (IET). *Journal of Military Ethics*, 9(1), 41–56. Available at: <https://doi.org/10.1080/15027570903523107>.
- Xu, M., Fralick, D., Zheng, J. Z., Wang, B., Tu, X. M., & Feng, C. (2017). The differences and similarities between two-sample T-test and paired T-test. *Shanghai Jingshen Yixue*, 29(3), 184–188.
- Yanikoski, R. (2004). Leadership perspectives on the role of character development in higher education. *New Directions for Institutional Research*, 122, 7–23.
- Zurqoni, Z., Retnawati, H., Arlinwibowo, J., & Apino, E. (2018). Strategy and implementation of character education in senior high schools and vocational high schools. *Journal of Social Studies Education Research*, 9(3), 370–397.

Views and opinions expressed in this article are the views and opinions of the author(s), International Journal of Education and Practice shall not be responsible or answerable for any loss, damage or liability etc. caused in relation to/arising out of the use of the content.