Journal of Education and Practice
ISSN 2222-1735 (Paper) ISSN 2222-288X (Online)
Vol.13, No.14, 2022

Effects of Peer Tutoring on Preservice Teachers' Performance on Practice Content Exams in South Texas, USA

Marsha Sowell
Teacher and Bilingual Education Department, Texas A&M University - Kingsville
700 University Blvd., MSC 196, Kingsville, Texas, 78363 USA
Tel: 1-361-593-4782 E-mail: marsha.sowell@tamuk.edu

The research is financed by the Southwest Educational Research Association Early Career Faculty Mentoring Initiative Grant.

Abstract
To become a teacher in the United States, teaching candidates must demonstrate understanding of good teaching practices and appropriate content knowledge in the subject to be taught (Kowarski, 2021). Teacher preparation programs are trying to balance TCs needs to learn material, engage in quality field experiences, and perform on high-stakes tests assessing content and pedagogy. This quasi-experimental study explored the impact of various interventions on teacher candidate's performance on practice content exams. The participants of the study included 30 elementary teacher candidates and 7 secondary teacher candidates enrolled in a teacher preparation program in south Texas, USA. Elementary participants were divided into a control group (N=13) using only the online modules and an experimental group (N=17) using online modules and receiving tutoring as interventions between tests. Secondary participants' interventions consisted of preparing and presenting content-specific tutoring for elementary candidates. Data were collected through pretest scores prior to intervention and posttest scores after intervention. Findings show a statistically significant difference in pre-post exam scores in all groups, but no statistical difference between posttest scores in the elementary control and experimental groups. It was concluded that all intervention methods between tests were effective however the various methods of intervention did not provide statistical differences in outcomes. This research was supported by the SERA Early Career Faculty Mentoring Initiative Grant.

Keywords: teacher candidate, high-stakes tests, test preparation, teacher preparation
DOI: 10.7176/JEP/13-14-05
Publication date: May 31st 2022

1. Introduction

The teaching profession has come under scrutiny as of late in regards to teacher preparation. As such, there is a growing movement to include high-stakes assessments to the traditional requirements to become a teacher in the United States: earn a bachelor's degree and complete an educator preparation program (Rao et al., 2021). These assessments include both an assessment to demonstrate understanding of good teaching practices and an assessment to demonstrate appropriate content knowledge in the subject to be taught (Kowarski, 2021). Teacher preparation programs are trying to balance teacher candidates' (TCs) needs to learn material, engage in quality field experiences, and perform on high-stakes tests.

To become a teacher in Texas, the content knowledge exam that must be passed is the Texas Examination of Educator Standards (TExES). For those desiring to teach elementary, the content knowledge exam, known as "Core Subjects EC-6", consists of five core contents: (1) English language arts and reading, (2) mathematics, (3) social studies, (4) science, and (5) fine arts including art, music, physical education, health, and theatre. Traditional teacher preparation programs do not include learning this content, but rather learning strategies to teach this content. Furthermore, elementary TCs do not usually enter the profession with a love of all subjects but rather a love for teaching. As such, elementary TCs sometimes struggle with demonstrating mastery of this wide range of content knowledge. Secondary TCs, by contrast, specialize in mastering content in one area and thus do not struggle with demonstrating mastery of their content on their exams despite a focus on methods rather than content within their teacher preparation programs.

Instructors at one south Texas university decided to utilize secondary TCs content expertise to provide elementary TCs with content knowledge support. To support elementary TCs' mastery of content knowledge, this university piloted an intervention plan in which secondary TCs specializing in various core contents tutored elementary TCs through the use of face-to-face tutoring, virtual tutoring via a social media platform, and material available via an online test preparation platform.

Two main questions guided this study:

1. Did interventions (face-to-face and virtual tutoring, use of online test preparation modules) influence performance on practice content exams for elementary teacher candidates?
2. Did planning and providing content-specific peer tutoring influence performance on practice content exams for secondary teacher candidates?

2.1 High-stakes test preparation for teacher candidates

Rao et al. (2021) explored TC's perspectives of their preparation program's supports of edTPA, a high-stakes assessment for evaluating TC's readiness to teach. Their study found that TCs reported strategic supports for the high-stakes assessments were helpful in passing the assessment but hindered the authenticity of the clinical experiences and the development of necessary teaching skills (Rao et al., 2021). In a similar study, Auslander et al. (2020) also found that TCs "were prioritizing the navigation of edTPA processes and requirements over learning how to become reflective and skilled teachers" (p. 384). These studies highlighted a disconnect between coursework including assessment preparation and authentic, practical learning experiences for TCs (Rao et al., 2021). In an effort to support the learning of high-stakes content knowledge while connecting with educational coursework and practical teaching experiences, the current study integrated content tutoring into a course on teaching methods. Secondary TCs used the cyclical lesson process to provide tutoring for elementary TCs, thus employing authentic teaching practices. Elementary TCs not only learned content from secondary TCs, but also analyzed and reflected on the content-specific teaching methods modeled by the secondary TCs.

Maddox and Reglin (2019) used a correlational design to explore teacher preparation practice tests as a predictor of TC success on teacher licensure exams in Virginia. They concluded that when practice tests were aligned in content to the real high-stakes test, there was a statistically significant positive correlation between practice exam performance and real exam performance. Furthermore, the positive correlations were moderate to large strength (Maddox and Reglin, 2019). Chou (2019) examined how multiple social factors influence self-efficacy for high-school seniors in preparing for high-stakes exams. This study concluded that while prior test performance greatly influenced self-efficacy in test preparation, self-efficacy could be increased through the use of strategies and resources in test preparation rather than test-taking practice alone (Chou, 2019). The current study sought to increase TCs content knowledge, thus ability to succeed on a teacher preparation practice test, by providing specific content knowledge strategies and resources. With stronger content specific strategies and knowledge, it was thought that TCs would perform better on teacher preparation practice tests and thus also on the real exams.

2.2 Peer tutoring

Peer tutoring is recognized as an effective intervention, improving academic performance for both tutors and tutees around the world (Campbell, 2019; Elbulok-Charcape et al., 2019; Omotayo & Lawrence, 2020; Petancio et al., 2020). Petancio et al., (2020) explain, "Tutees can benefit by getting help from a peer in clarifying a concept; tutors can benefit as they strengthen their understanding of a concept by answering the tutees' questions" (p. 49). Omotayo and Lawrence (2020) add that teacher candidates earn double benefits from peer tutoring. As they provide tutoring, they experience preparing and presenting instruction in their content area. Furthermore, both tutors and tutees participate in an "innovative teaching strategy in the course of their training and find it more convenient to apply when they are in the field practicing as teachers" (p. 23). This study involves secondary TCs acting as tutors for elementary TCs within the context of their teacher training.

Elbulok-Charcape et al. (2019) found a positive association between the frequency of peer tutoring and improvement on exam scores. As such, they recommend instructors make attending tutoring a course requirement. Face-to-face tutoring, however, can be costly in both time and money. Campbell (2019) reports combined face-to-face and virtual tutoring increased student engagement in tutoring, but virtual tutoring alone can be effective while saving both time and money. Crowley-Cyr and Hever's (2021) studied law students participating in online peer assistance programs with no face-to-face component. Participants in their study performed better on exams, increased engagement in group forums, and were more consistent in accessing online study materials during key study periods than students not participating in the online peer assistance programs. Petancio et al. (2020) summarized "in the new normal of education, peer tutoring then still has a place with online delivery modality" (p. 49). This study involves a group of elementary teacher candidates participating in combined face-to-face and virtual tutoring interventions.

Petancio et al. (2020) note that peer tutoring in higher education improves academic performance but few studies explore the impact of peer tutoring on subject mastery. Furthermore, they recommend future research in varied subject areas. Campbell (2019) recognizes that despite the increased use of social media for educational purposes, there is little research in utilizing social media as a tutoring platform. The proposed study will use pre-post test analysis to determine elementary TCs' content mastery in multiple subjects through the use of both face-to-face and virtual tutoring on a social media platform.

The purpose of this study is to explore the impact of various test preparation interventions on TCs' demonstration of content knowledge mastery through performance on content preparation exams.

3. Method

3.1 Data collection instrument

This quasi-experimental study utilized repeated practice exams through Certify Teacher

(www.certifyteacher.com), an online test preparation software for educator certification. Certify Teacher offers multiple full-length practice exams aligned to current test standards, thus ensuring that each practice exam offers different questions rather than repeated questions (Simulados Software, Inc, 2022). This study utilized the Texas Examination of Educator Standards (TExES) practice exams for various content areas.

Each TExES practice exam is aligned to the content and format of the current test standards. The Core Subjects EC-6 exam consists of 230 selected-response questions in five domains: English Language Arts and Reading & Science of Teaching Reading (ELAR & STR), Mathematics, Social Studies, Science, and Fine Arts (Pearson Education, 2022). All elementary TCs took this exam. Secondary TCs took the exams that correlated with their future teaching content area, including: English Language Arts and Reading 7-12, History 7-12, Life Science 7-12, Mathematics 7-12, Music EC-12, and Technology Applications EC-12. The history, life science, mathematics, music, and technology applications exams each consist of 100 selected-response questions. The English Language Arts and Reading exam consists of 100 selected response and two constructed response questions. The social studies exam consists of 140 selected response questions (Pearson Education, 2022).

3.2 Intervention measures

In this quasi-experimental study, elementary TCs were divided into experimental (N=17) and control (N=13) groups based on their instructor for a selected education methods course. Each group took a TExES Core Subjects EC-6 (early childhood through sixth grade) pretest through the online test preparation platform in August 2021 and continued using the platform to study for this exam. In addition to the online platform, elementary TCs in the experimental group received one face-to-face tutoring session in September 2021 and three virtual tutoring sessions throughout September and October 2021. The tutoring sessions were hosted by secondary TCs and included tutoring in each of the five domains of the Core Subjects EC-6 exam. Teacher candidates in the control group did not receive additional tutoring. All elementary TCs then took a TExES Core Subjects EC-6 posttest through the online test preparation platform in October 2021.

Secondary TCs took a TExES pretest through the online test preparation platform in August 2021. Each of these TCs took the exam that aligned to their selected content area (i.e. a future history teacher took the History 7-12 exam). Through classwork, secondary TCs analyzed elementary preservice teacher's pretest scores, identified areas in need of tutoring, and prepared and presented one face-to-face and three virtual tutoring sessions for elementary TCs. Each secondary TC focused their data analysis and tutoring on their content area (i.e. a future history teacher focused on data and tutoring in social studies). In October 2021, secondary TCs then took a TExES posttest in their selected content area through the online test preparation platform. Although 14 secondary TCs participated in the tutoring assignments, several had already passed their official TExES content exam and did not utilize the online platform, thus only the data from seven secondary TCs was included in this study.

4. Results

4.1 Normality of difference in pretest and posttest practice exam scores

Data was analyzed using IBM SPSS Statistics version 27. To determine the appropriate statistical method to use in comparing the pretest and posttest scores of each group, and due to the small sample size, a Shipiro-Wilk Test of Normality was performed on the differences in overall test scores as well as differences in scores for each domain within the Core Subjects EC-6 exam.

As seen in Table 1, the Shapiro-Wilk Test indicated the difference in scores for elementary TC participants in the control group were normally distributed for the overall exam scores (W =.93, p-value = .312), ELAR & STR (W = .88, p-value = .061), mathematics (W = .93, p-value = .315), social studies (W = .90, p-value = .114), science (W = .91, p-value = .143), and fine arts (W = .95, p-value = .143). These results indicated a paired-samples t-test would be appropriate for further analysis in all areas for the control group.

Further noted in Table 1, the Shapiro-Wilk test indicated the difference in scores for elementary TC participants in the experimental group were normally distributed for the overall exam scores (W =.92, p-value = .123), ELAR & STR (W = .94, p-value = .306), mathematics (W = .94, p-value = .346), social studies (W = .91, p-value = .095), and science (W = .95, p-value = .379). These results indicated a paired-samples t-test would also be appropriate for further analysis of the experimental groups' scores for the overall exam, ELAR & STR, mathematics, social studies, and science. Fine arts scores for the experimental group, however, departed significantly from normality (W = .85, p-value = .010), thus indicating the nonparametric Related-Samples Wilcoxon Signed Rank Test would be more appropriate for pretest and posttest comparisons in this domain.

The normality of pretest and posttest score differences was also assessed for the secondary content area exams. The Shapiro-Wilk Test of Normality did not show evidence of non-normality (W = .893(7), p-value =.291) for secondary scores. Based on this outcome, the paired-samples t-test was deemed appropriate to compare the secondary pretest and posttest scores.

4.2 Comparing pretest and posttest scores within the control group

To compare the pretest and posttest scores within the control group, a paired-samples *t*-test was conducted on the control group's scores for the overall practice TExES EC-6 Core Subjects exam as well as for scores in each of the five domains of the exam. Elementary TCs in the control group were able to access the online test preparation software modules to study between the pretest and posttests, but did not receive any formal additional tutoring between the administration of the exams.

As shown in Table 2, the *t*-test indicated scores for the control group (*N =13)* were associated with an increased score in each area analyzed. The overall score was associated with an increased score $M = 16.50$ ($SD = 11.70$). The ELAR & STR domain score was associated with an increased score $M = 15.35$ ($SD = 18.31$). The mathematics domain score was associated with an increased score $M = 17.17$ ($SD = 14.26$). The social studies domain score was associated with an increased score $M = 21.76$ ($SD = 17.75$). The science domain score was associated with an increased score $M = 16.83$ ($SD = 15.08$). The fine arts domain score was associated with an increased score $M = 12.37$ ($SD = 14.70$).

For the control group, the null hypothesis of equal scores in the overall exam was rejected, $t(12) = 5.09$, $p < .000$. Likewise, the null hypothesis of equal scores in the ELAR & STR domain was rejected, $t(12) = 3.02$, $p = .011$, as was the null hypothesis of equal scores in the mathematics domain, $t(12) = 4.34$, $p = .001$. The null hypothesis of equal scores in the social studies domain was also rejected, $t(12) = 4.42$, $p = .001$ as was the null hypothesis of equal scores in the science domain, $t(12) = 4.02$, $p = .002$. The null hypothesis of equal scores in the fine arts domain was also rejected, $t(12) = 3.03$, $p = .010$ (refer to Table 2). Thus, the posttest means of each area were statistically significantly higher than the pretest means of each area.

Cohen's *d* was estimated at 1.41 for the overall exam scores, .84 for the ELAR & STR scores, 1.20 for the mathematics scores, 1.23 for the social studies scores, 1.17 for the science scores, and .84 for the fine arts scores. The estimations for the overall exam scores exceed Cohen's (1988) conventions for a very large effect while each of the domains exceed Cohen's conventions for a large effect (refer to Table 2).

4.3 Comparing pretest and posttest scores within the experimental group

To compare the pretest and posttest scores within the experimental group, a paired-samples *t*-test was conducted on the group's scores for the overall practice TExES EC-6 Core Subjects exam as well as for scores in the ELAR & STR, mathematics, social studies, and science domains. Due to non-normality of scores, the Related-Samples Wilcoxon Signed Rank test was conducted for fine arts domain of the experimental group. Elementary TCs in the experimental group were able to access the online test preparation software modules to study between the pretest and posttests, but also received formal tutoring between the administration of the exams. Formal tutoring included one face-to-face tutoring session and three virtual tutoring sessions conducted by secondary TCs.

As shown in Table 3, the *t*-test indicated scores for the experimental group (*N =17)* were associated with an increased score in each area analyzed. The overall score was associated with an increased score $M = 9.71$ ($SD = 11.99$). The ELAR & STR domain score was associated with an increased score $M = 9.97$ ($SD = 16.37$). The mathematics domain score was associated with an increased score $M = 3.00$ ($SD = 8.27$). The social studies domain score was associated with an increased score $M = 12.94$ ($SD = 21.07$). The science domain score was associated with an increased score $M = 9.77$ ($SD = 15.80$). As shown in Table 4, the fine arts domain score was associated with an increased score $M = 13.06$.

For the experimental group, the null hypothesis of equal scores in the overall exam was rejected, $t(16) = 3.34$, $p = .004$. Likewise, the null hypothesis of equal scores in the ELAR & STR domain was rejected, $t(16) = 2.51$, $p = .023$, as was the null hypothesis of equal scores in the social studies domain, $t(16) = 2.53$, $p = .022$. The null hypothesis of equal scores in the science domain was also rejected, $t(16) = 2.55$, $p = .021$ (refer to Table 3). As shown in Table 4, the Related-Samples Wilcoxon Signed Rank Test indicated there was a significant difference in the scores for the pretest and posttest scores for fine arts for the experimental group as well *($z = 2.627$ $p = .009$)*. Only the mathematics domain scores within the experimental group did not show a significant difference between the pretest (*M=74.05. D = 13.88*) and posttest (*M = 77.04, SD = 15.79*); $t(16) = 1.49$, $p = .155$; thus the null hypothesis was retained for the experimental group's mathematical domain scores (refer to Table 3). The posttest means were statistically significantly higher than the pretest means for the experimental group's overall exam scores, ELAR & STR, social studies, science, and fine arts.

Within the experimental group, effect sizes varied. The overall test score ($d = .81$) and fine arts ($r = .64$) indicated a large effect based on Cohen's (1988) guidelines. Effect sizes for the experimental group's ELAR & STR domain ($d = .60$), social studies domain ($d = .61$), and science domain ($d = .62$) were found to exceed Cohen's conventions for medium effect. The mathematics domain scores ($d = .36$), although not demonstrating statistical significance, did exceeded Cohen's convention for small effect (refer to Table 3 and Table 4).

4.4 Comparing pretest and posttest scores within the secondary preservice teacher group

A paired-samples *t*-test was conducted to compare the pretest and posttest scores for secondary TCs taking

various content practice exams. There was a significant difference in the pretest scores (*M=66.14, SD=17.85*) and the posttest scores (*M=88.50, SD=8.56*); *t(6)=3.78, p=.009*, thus the null hypothesis of equal scores was rejected. The effect size analysis for secondary content tests (*d=1.43*) was found to exceed Cohen's (1988) convention for a very large effect.

4.5 Normality of means in pretest and posttest TExES EC-6 Core Subjects practice exam scores

To determine the appropriate statistical method to use in comparing the practice Core Subjects EC-6 exam scores between the control and experimental groups, and due to the small sample size, a Shipiro-Wilk Test of Normality was performed on the means of pretest and posttest scores for the overall exam and each domain within the exam.

As seen in Table 5, for elementary TC participants in the control group, the Shapiro-Wilk tests indicated the pretest scores were normally distributed for the overall scores (*W =.95, p-value = .658*), ELAR & STR (*W =.90, p-value = .127*), mathematics (*W =.92, p-value = .244*), social studies (*W =.96, p-value = .679*), science (*W =.94, p-value = .487*), and fine arts (*W =.95, p-value = .557*). For elementary TC participants in the experimental group, the Shapiro-Wilk tests indicated the pretest scores were normally distributed for the overall scores (*W =.94, p-value = .373*), ELAR & STR (*W =.95, p-value = .498*), mathematics (*W =.94, p-value = .344*), social studies (*W =.94, p-value = .302*), and science (*W =.90, p-value = .076*), but not for fine arts (*W =.83, p-value = .005*). Based on this outcome, the independent samples *t*-test was deemed appropriate to compare the pretest scores for the two participant groups in the overall exam, ELAR & STR, mathematics, social studies, and science. The use of the Mann-Whitney U Test was deemed appropriate to compare the pretest scores for the two participant groups in the fine arts domain.

As seen in Table 6, for elementary TC participants in the control group, the Shapiro-Wilk tests indicated the posttest scores were normally distributed for the science domain (W = .92, *p*-value = .268), but not for the overall scores (*W =.83, p-value = .016*), ELAR & STR (*W =.81, p-value = .009*), mathematics (*W =.82, p-value = .010*), social studies (*W =.82, p-value = .011*), or fine arts (*W =.80, p-value = .007*). For elementary TC participants in the experimental group, the Shapiro-Wilk tests indicated the posttest scores were normally distributed for the mathematics domain (*W =.93, p-value = .195*), but not for the overall scores (*W*= .86, *p-value* = .017), ELAR & STR (*W =.85, p-value = .010*), social studies (*W =.89, p-value = .040*), science (*W =.83, p-value = .005*), or fine arts (*W =.89, p-value = .046*). These results indicated the use of the Mann-Whitney U Test would be appropriate for comparing posttest scores in all areas.

4.6 Comparing practice pretest TExES Core Subjects EC-6 exam scores between groups

As seen in Table 5, the control group (*N* = 13) was associated with an overall pretest score *M* = 62.73 (*SD* = 18.13). By comparison, the experimental group *(N* = 17) was associated with a higher overall pretest score *M* = 67.50 (*SD* = 13.54). The control group was associated with an ELAR & STR pretest score *M* = 67.75 (*SD* = 22.03). By comparison, the experimental group was associated with a higher ELAR & STR pretest score *M* = 69.83 (*SD* = 19.47). The control group was associated with a mathematics pretest score *M* = 65.99 (*SD* = 18.02). By comparison, the experimental group was associated with a higher mathematics pretest score *M* = 74.05 (*SD* = 13.88). The control group was associated with a social studies pretest score *M* = 59.12 (*SD* = 22.75). By comparison, the experimental group was associated with a higher social studies pretest score *M* = 62.86 (*SD* = 16.84). The control group was associated with a science pretest score *M* = 57.25 (*SD* = 20.06). By comparison, the experimental group was associated with a higher science pretest score *M* = 66.21 (*SD* = 15.42). The control group was associated with a fine arts pretest score *M* = 62.67 (*SD* = 19.89). By comparison, the experimental group was associated with a higher fine arts pretest score *M* = 64.23 (*SD* = 20.56).

As seen in Table 7, the overall exam scores were not associated with a statistically significant effect (*t* = .83, *p* = .416). The ELAR & STR domain scores were not associated with a statistically significant effect (*t* = 27, *p* = .786). The mathematics scores were not associated with a statistically significant effect (*t* = 1.39, *p* = .177). The social studies scores were not associated with a statistically significant effect (*t* = .52, *p* = .608). The science scores were not associated with a statistically significant effect (*t* = 1.39, *p* = .177). As seen in Table 8, the fine arts scores were not associated with a statistically significant effect (*U* = 102.5, *p* = .738). Thus, while the pretest scores of the experimental group were higher they were not found to be statistically significantly higher than the pretest scores of the control group, thus the null hypothesis of equal scores is retained in all areas.

Cohen's *d* was estimated for the mathematic domain score at .51 and for the science domain score at .51 which indicates a medium effect based on Cohen's (1988) guidelines. Effect sizes for the overall exam (*d* = .30) was found to exceed Cohen's conventions for small effect. The ELAR & STR domain (*d* = .10), social studies domain (*d* = .19), nor fine arts domain (*r* = .06) met the conventions set forth for small effect. (refer to Table 7 and Table 8).

4.7 Comparing practice posttest TExES Core Subjects EC-6 exam scores between groups

As seen in Table 6, the control group (*N* = 13) was associated with an overall posttest score *M* = 79.23 (*SD* =

17.76). By comparison, the experimental group *(N* = 17) was associated with a lower overall posttest score *M* = 77.21 (*SD* = 15.20). The control group was associated with an ELAR & STR posttest score *M* = 83.10 (*SD* = 16.81). By comparison, the experimental group was associated with a lower ELAR & STR posttest score *M* = 79.80 (*SD* = 17.43). The control group was associated with a mathematics posttest score *M* = 83.16 (*SD* = 15.03). By comparison, the experimental group was associated with a lower mathematics posttest score *M* = 77.04 (*SD* = 15.79). The control group was associated with a social studies posttest score *M* = 80.88 (*SD* = 23.22). By comparison, the experimental group was associated with a lower social studies posttest score *M* = 76.00 (*SD* = 19.82). The control group was associated with a science posttest score *M* = 74.08 (*SD* = 16.20). By comparison, the experimental group was associated with a higher science posttest score *M* = 75.98 (*SD* = 18.03). The control group was associated with a fine arts posttest score *M* = 75.04 (*SD* = 26.90). By comparison, the experimental group was associated with a higher fine arts posttest score *M* = 77.28 (*SD* = 14.33).

As seen in Table 9, the overall exam scores were not associated with a statistically significant effect (*U* = 85.00, *p* = .285). The ELAR & STR domain scores were not associated with a statistically significant effect (*U* = 91.50, *p* = .426). The mathematics scores were not associated with a statistically significant effect (*U* = 85.50, *p* = .294). The social studies scores were not associated with a statistically significant effect (*U* = 84.00, *p* = .266). The science scores were not associated with a statistically significant effect (*U* = 101.50, *p* = .706). The fine arts scores were not associated with a statistically significant effect (*U* = 98.50, *p* = .615). Thus, while the posttest scores of the control group were higher than the experimental group in the overall exam, ELAR & STR, mathematics, and social studies, they were not found to be statistically significantly higher. Likewise, while the posttest scores of the experimental group were higher than the control group in the science and fine arts domains, they were not found to be statistically significantly higher. The null hypothesis of equal scores is retained in all areas for posttest scores.

The effect size analysis was found to meet Cohen's (1988) convention for a small effect in overall test scores (*r* = .19), ELAR & STR (*r* = .15), mathematics (*r* = .19), and social studies (*r* = .20). These are the areas in which the control group outperformed the experimental group on the posttest. Effect sizes in science (*r* = .07) and fine arts (*r* = .09), the areas in which the experimental group outperformed the control group, did not meet Cohen's conventions for a small effect (refer to Table 9).

5. Discussion and Conclusions

5.1 Did interventions (face-to-face and virtual tutoring, use of online test preparation modules) influence performance on practice content exams for elementary TCs?

Elementary TCs' scores on the practice TExES Core Subjects EC-6 exams demonstrated a statistically significant improvement between their pretest and posttest attempts, indicating that test preparation interventions were successful. While the interventions provided students were successful in improving test scores for elementary TCs, the method of intervention did not demonstrate a statistically significant difference in the improvement of scores.

While both the control and experimental groups demonstrated large effects for the overall exam score improvement, only the gains for the control group met Cohen's criteria for large effects in each domain. The experimental group only met Cohen's criteria for medium or small effects within most of the individual domains. This may indicate that these small gains work together to provide a large overall gain in exam. Because both groups had access to the online modules, this discrepancy in practical effectiveness may not be attributed to the intervention method. The posttest scores do not show a statistical difference between the two groups, indicating the two groups achieved similar posttest scores. While the pretest means were not statistically different between the two groups, the means of the experimental pretests were higher than the means of the control group. This could indicate that the experimental group needed to demonstrate less growth to achieve the expected score for moving on in the educational program than the control group needed to demonstrate, but the interventions met the needs of both groups.

5.2 Did planning and providing content-specific peer tutoring influence performance on practice content exams for secondary TCs?

Like the elementary TCs' scores, the secondary TCs demonstrated significant improvement in their practice TExES content exams, indicating that test preparation interventions were also successful for secondary TCs. Secondary teachers do not take the same test and thus scores cannot be evaluated by domain, their overall score improvements were found to exceed Cohen's convention for very large effects. Like elementary teachers, this effect could be due to the need for large improvements to meet the expected score for moving on in their educational program.

The results of this study indicate that exam preparation interventions are successful in improving preservice teacher's practice exam scores. While effect sizes of intervention vary, this may be due to the student's individual improvement needed to meet a given standard and thus their motivation to improve in the test.

Focused intervention strategies should be planned and implemented in teacher preparation programs to ensure TCs are prepared to demonstrate content knowledge on high-stakes tests (Chou, 2019).

6. Limitations and Recommendations for Future Research

While this study indicates test preparation interventions are effective in improving TCs' performance on practice high-stakes exams, it fails to indicate the types of interventions that are most effective. It is recommended that future studies separate each form of intervention to measure direct influence. Furthermore, this pilot study is limited in the number of participants and should be expanded to additional courses, universities, and types of high-stakes content exams for teachers. It is also believed that the TCs benefited from the peer tutoring in ways that may not be demonstrated on an online exam (Campbell, 2019; Elbulok-Charcape et al., 2019; Omotayo & Lawrence, 2020; Petancio et al., 2020) and these other benefits should be explored in future studies as well.

References

Auslander, S.S., Myers, B., Tanguay, C., Smith, S.Z., Myers, K.D. (2021). High-stakes assessment in an elementary teacher preparation program: A case study of multiple stakeholders. *Teacher Development, 25*(3), 366-388. https://doi.org/10.1080/13664530.2021.1915371

Campbell, A. (2019). Design-based research principles for successful peer tutoring on social media. *International Journal of Mathematical Education in Science and Technology, 50*(7), 1024-1036. http://dx.doi.org/10.1080/0020739X.2019.1650306

Chou, M-H. (2019). Predicting self-efficacy in test preparation: Gender, value, anxiety, test performance, and strategies. *The Journal of Educational Research, 112*(1), 61-71. https://doi.org/10.1080/00220671.2018.1437530

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Crowley-Cyr, L., & Hevers, J. (2021). Using peer assisted learning to improve academic engagement and progression of first-year online law students. *Journal of University Teaching & Learning Practice, 18*(1), 1-17. https://doi.org/10.53761/1.18.1.2

Elbulok-Charcape, M., Grandoit, E., Berman, L., Fogel, J., Fink, L., & Rabin, L. (2019). Improving exam performance in an undergraduate statistics course for at-risk students through peer tutoring. *Psychology Teaching Review, 25*(2), 3-17. https://doi.org/10.1177/1475725719886399

Kowarski, I., (2021, July 1). How to become a licensed or certified teacher. *US News and World Report.* https://www.usnews.com/education/best-graduate-schools/articles/how-to-get-a-teaching-degree-and-become-a-teacher

Maddox, A., & Reglin, G. (2019). Teacher preparation tests and grade point average as a predictor of teacher licensure high-stakes tests. *College Student Journal, 52*(2), 229-242. https://www.ingentaconnect.com/contentone/prin/csj/2019/00000053/00000002/art00009

Omotayo, K.A. & Lawrence, A.D. (2020). Effect of peer-tutoring on pre-service teachers' performance in basic science in Nigerian Colleges of Education. *Commonwealth Journal of Academic Research, 1*(3), 23-31.

Pearson Education. (2022). *Preparation manuals.* Texas Educator Certification Examination Program. http://www.tx.nesinc.com/PageView.aspx?f=HTML_FRAG/GENRB_PrepManuals.html

Petancio, J.M., Perez, N.B., & Javier, N.G.N. (2020). Peer tutoring: Its effects on subject mastery and mathematics anxiety among elementary teaching interns. *Jurnal Pendidkan Malaysia, Vol 45*(2), 46-57. http://dx.doi.org/10.17576/JPEN-2020-45.02-05

Rao, A.B., Olson, J.D., & Koss, M.D. (2021). Teacher candidate perspectives of edTPA readiness: Preparation programs and edTPA supports. *Teaching Education, 32*(3), 251-268. http://dx.doi.org/10.1080/10476210.2020.1713079

Simulados Software, Inc. (2022). *List of products, Certify Teacher.* Certify Teacher: Practice Software for Educator Certification. https://www.certifyteacher.com/produtos/list/texas

Tables

Table 1: Normality of practice TExES Core Subjects EC-6 pretest and posttest score differences

Exam Component	Control Group		Experimental Group	
	Statistic(13)	Sig.	Statistic(17)	Sig.
Overall	.93	.312	.92	.123
ELAR & STR	.88	.061	.94	.306
Mathematics	.93	.315	.94	.346
Social studies	.90	.114	.91	.095
Science	.90	.143	.95	.379
Fine art	.95	.518	.85	.010*

*$p < .05$

Shapiro-Wilks Test of Normality results for TExES Core Subjects EC-6 pretest and posttest score differences

Table 2: Comparing control group pretest and posttest scores

	Mean Difference (posttest-pretest)	Std Dev	Paired *t*-test		
			t(12)	Sig (2-tailed)	Cohen's *d*
Overall	16.5	11.70	5.09	.000	1.41
ELAR & STR	15.35	18.31	3.02	.011	.84
Mathematics	17.17	14.26	4.34	.001	1.20
Social studies	21.76	17.75	4.42	.001	1.23
Science	16.82	15.08	4.02	.002	1.17
Fine arts	12.37	14.70	3.03	.010	.84

*$p > .05$

Paired sample *t*-test results comparing the control group's pretest and posttest practice TExES Core Subjects EC-6 scores.

Table 3: Comparing experimental group pretest and posttest for normally distributed scores

	Mean Difference (posttest-pretest)	Std Dev	Paired *t*-test		
			t (16)	Sig (2-tailed)	Cohen's *d*
Overall	9.71	11.99	3.34	.004	.81
ELAR & STR	9.97	16.37	2.51	.023	.61
Mathematics	3.00	8.27	1.49	.155*	.36
Social studies	12.94	21.07	2.53	.022	.61
Science	9.77	15.80	2.55	.021	.62

*$p > .05$

Paired sample *t*-test results comparing the experimental group's pretest and posttest practice TExES Core Subjects EC-6 scores overall, and in the ELAR & STR, mathematics, social studies, and science domains.

Table 4: Comparing experimental group pretest and posttest for non-normally distributed scores

	Mean Difference (posttest-pretest)	Std Dev	Wilcoxon Signed Ranks Test		
			Z	*p*	*r*
Fine Arts	13.06	11.99	2.63	.009	.64

*$p > .05$

Wilcoxon Signed Ranks Test results comparing the experimental group's pretest and posttest practice TExES Core Subjects EC-6 scores for the fine arts domain

Table 5: Normality of practice TExES Core Subjects EC-6 pretest mean scores

Exam Component	Control Group				Experimental Group			
	Mean	Std Dev	Statistic (13)	Sig.	Mean	Std Dev	Statistic (17)	Sig.
Overall	62.73	18.13	.95	.658	67.50	13.54	.94	.373
ELAR & STR	67.75	22.03	.90	.127	69.83	19.47	.95	.498
Mathematics	65.99	18.02	.92	.244	74.05	13.88	.94	.344
Social Studies	59.12	22.75	.96	.679	62.86	16.84	.94	.302
Science	57.25	20.06	.94	.487	66.21	15.42	.90	.076
Fine Arts	62.67	19.89	.95	.557	64.23	20.56	.83	.005*

*$p < .05$

Shapiro-Wilks Test of Normality results for TExES Core Subjects EC-6 pretest mean scores

Table 6: Normality of practice TExES Core Subjects EC-6 posttest mean scores

Exam Component	Control Group				Experimental Group			
	Mean	Std Dev	Statistic(13)	Sig.	Mean	Std Dev	Statistic (17)	Sig.
Overall	79.23	17.76	.83	.016*	77.21	15.20	.86	.017*
ELAR & STR	83.10	16.81	.81	.009*	79.80	17.43	.85	.010*
Mathematics	83.16	15.03	.82	.010*	77.04	15.79	.93	.195
Social Studies	80.88	23.22	.82	.011*	76.00	19.82	.89	.040*
Science	74.08	16.20	.92	.268	75.98	18.03	.83	.005*
Fine Arts	75.04	26.90	.80	.007*	77.28	14.33	.89	.046*

*$p < .05$
Shapiro-Wilks Test of Normality results for TExES Core Subjects EC-6 posttest mean scores

Table 7: Comparison of practice TExES Core Subjects EC-6 pretest normally distributed scores

	Levene's *F* test		Independent-samples *t*-test		
Exam Component	*F*	Sig.	*t(28)*	Sig (2-tailed)	*Cohen's d*
Overall	1.37	.251*	.83	.416*	.30
ELAR & STR	.57	.457*	.27	.786*	.10
Mathematics	.32	.576*	1.39	.177*	.51
Social Studies	.96	.336*	.52	.608*	.19
Science	2.09	.160*	1.39	.177*	.51

*$p > .05$
Independent samples *t*-test comparing control group and experimental group practice TExES Core Subjects EC-6 exam pretest scores overall and in ELAR & STR, mathematics, social studies, and science domains.

Table 8: Comparison of practice TExES Core Subjects EC-6 pretest non-normally distributed scores

	Median Scores		Mann-Whitney Test			
Exam Component	*Control*	*Experimental*	*U* statistic	*z-score*	*p*	*r*
Fine Arts	63.41	66.67	102.5	.335	.738	.06

Note: Control (N=13), Experimental (N=17)
Mann-Whitney Test comparing control group and experimental group practice TExES Core Subjects EC-6 exam pretest scores for the fine arts domain.

Table 9: Comparison of practice TExES Core Subjects EC-6 posttest scores

	Median Scores		Mann-Whitney Test			
Exam Component	*Control*	*Experimental*	*U* statistic	*z-score*	*p*	*r*
Overall	92.00	84.50	85.00	-1.09	.285	.19
ELAR & STR	93.85	87.69	91.50	-.80	.426	.15
Mathematics	90.00	76.19	85.50	-1.05	.294	.19
Social Studies	88.57	85.71	84.00	-1.11	.266	.20
Science	80.85	82.61	101.50	-.38	.706	.07
Fine Arts	90.24	83.72	98.50	-.50	.615	.09

Note: Control (N=13), Experimental (N=17)
Mann-Whitney Test comparing control and experimental group practice TExES Core Subjects EC-6 exam posttest scores for overall, ELAR & STR, mathematics, social studies, science, and fine arts domains.