10-18-2018

Life Science Undergraduate Mentors in NE STEM 4U Significantly Outperform Their Peers in Critical Thinking Skills

Kari L. Nelson

Claudia M. Rauter

Christine E. Cutucache

Life Science Undergraduate Mentors in NE STEM 4U Significantly Outperform Their Peers in Critical Thinking Skills

Kari L. Nelson, Claudia M. Rauter, and Christine E. Cutucache*
Department of Biology, University of Nebraska at Omaha, Omaha, NE 68182-0040

ABSTRACT
The development of critical thinking skills in recent college graduates is keenly requested by employers year after year. Moreover, improving these skills can help students to better question and analyze data. Consequently, we aimed to implement a training program that would add to the critical thinking skills of undergraduate students: Nebraska Science, Technology, Engineering, and Math 4U (NE STEM 4U). In this program, undergraduates provide outreach, mentoring, and science, technology, engineering, and mathematics (STEM) education to K–8 students. To determine the impacts of serving as an undergraduate mentor in this program on critical thinking, we compared undergraduate mentors (intervention group) with nonmentor STEM majors (nonintervention, matched group) using the valid and reliable California Critical Thinking Skills Test (CCTST) as a pre/post measurement. Importantly, before the intervention, both NE STEM 4U mentors and nonmentor undergraduates scored similarly overall on the CCTST. However, the posttest, carried out one academic year later, indicated significant gains in critical thinking by the NE STEM 4U mentors compared with the nonmentors. Specifically, the math-related skills of analysis, inference, and numeracy improved significantly in mentors compared with nonmentors.

INTRODUCTION
Critical thinking is a skill routinely cited as preferred by employers over basic content understanding (National Science Board, 2010; Association of American Colleges and Universities, 2013; National Association of Colleges and Employers [NACE], 2014; New York Academy of Sciences, 2014; National Academies of Sciences, Engineering, and Medicine, 2016) and is a core learning objective of science education (Phillips and Bond, 2004; Carneval *et al*., 2010; Langdon *et al.*, 2011; Dowd *et al.*, 2018). Moreover, as the employment landscape becomes more competitive, it is imperative that students have the opportunity for a dynamic, well-rounded professional development experience at the college level. The acquisition of so called “soft skills” such as critical thinking translates across areas of content expertise, not excluding the sciences. Undergraduates (UGs) are increasingly being encouraged to become involved in activities such as mentoring and service learning to develop these soft skills. While there is demonstrated impact of mentoring and service learning on the mentees, there is little empirical evidence of the impacts of these activities on the UG mentors themselves (Carpenter, 2015; Nelson and Cutucache, 2017). Given this, we studied whether participating in the intervention of Nebraska Science, Technology, Engineering, and Math 4U (NE STEM 4U) as a UG mentor impacted the critical thinking skills of UG life science majors.

NE STEM 4U
The NE STEM 4U program is a professional training program for UG and graduate students in science, technology, engineering, and mathematics (STEM) majors who provide outreach using inquiry-based learning to students in grades K–8

Elisabeth Schussler, *Monitoring Editor*
Submitted Mar 14, 2018; Revised Jul 6, 2018; Accepted Jul 12, 2018
CBE Life Sci Educ December 1, 2018 17:ar54
DOI:DOI: 10.1187/cbe.18-03-0038
Conflict of interest disclosure: C.E.C. designed and created the NE STEM 4U intervention, and K.L.N. conducted all of the critical thinking assessments using the instrument straight from publisher without any raw data analysis or processing by C.E.C. All raw data were processed by K.L.N. and C.M.R. to avoid any potential bias from the intervention creator (i.e., C.E.C.).
*Address correspondence to: Christine E. Cutucache (ccutucache@unomaha.edu).

(Cutucache *et al.*, 2016). NE STEM 4U, as a program, utilizes a threefold training platform of teaching, mentoring, and research. Specifically, all participants must serve as teachers in the program and provide mentorship to youth, and all participants receive mentorship from faculty advisors and peers throughout the program.

The UG mentors in the program are students in STEM majors (predominantly from the life sciences) who have little or no background in working with youth, after-school programming, research, teaching, or any formal mentorship. Prospective mentors apply to the program and go through an interview process and a background check to work with youth. After a student is formally admitted to the program, he or she shadows a veteran group of mentors two to three times before being partnered with a veteran mentor(s) to implement his or her own teaching, using inquiry-based practices (i.e., not traditional, transmittal lecture). Using a team-teaching model, we aim to have two to three NE STEM 4U mentors per school. In addition to these practices, there are once-monthly "experiment nights" and "STEMinars." At experiment nights, the mentors troubleshoot, in teams, the curriculum for that month. During STEMinars, we host speakers to present on a range of topics applicable to the out-of-school time teaching realm, such as classroom management, engaging special needs students, multicultural awareness, youth voice, relationship building, empathy, and understanding current budgetary challenges schools face. UG mentors also have the opportunity to participate in STEM education research.

Once the mentors are in the classroom, they engage the youth twice weekly at each school for 1 hour each time. Approximately half of our mentors mentor at more than one school per week. We estimate that the workload of 1 hour of teaching per week equates to a total of 4 hours invested for the mentor for teaching preparation and transportation to and from the site. In the classroom, the mentors engage no more than 15 young people at a time. To aid in relationship building, most schools ensure that the same K–8 students attend each time. The demographics of participating mentors/teachers is included in the *Methods* section and Supplemental Table 1. Of the students who have graduated from the program thus far ($n = 117$), 95.9% of them have both completed an academic degree in STEM and entered into the STEM workforce or graduate school upon graduation.

Critical Thinking

Critical thinking is delineated by a wide variety of definitions, although most agree with the philosophies of Socrates, Plato, and Aristotle (Gutek, 2009; Bailey and Mentz, 2015). Holyoak and Morrison (2005), describing critical thinking as an array of cognitive processes surrounding everyday life (Holyoak and Morrison, 2005; Dowd *et al.*, 2018). One of the most cited definitions of critical thinking comes from what has come to be called "the Delphi Report," in which 46 critical thinking experts across many disciplines came together to define critical thinking as "purposeful, self-regulatory judgment, which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which judgment is based" (Facione, 1990, p. 2). It is the Delphi Report that provides the foundation for the design of the California Critical Thinking Skills Test (CCTST), which has been used worldwide to measure critical thinking for more than 25 years (Insight Assessment, 2017) and was used in the current study.

Research on fostering critical thinking skills has primarily focused on *how* to teach critical thinking skills (Rowe *et al.*, 2015; Paris, 2016; Dowd *et al.*, 2018; Watanabe-Crockett, 2018). The majority of these studies investigated the influence of inquiry-based and problem-based teaching methods on thinking skills and showed that these instruction methods enhance the thinking skills of the learner (Greenwald and Quitadamo, 2014; Magrabi *et al.*, 2018). Research has also been conducted in the context of peer- and near-peer–led team learning. For example, Gellin (2003), Quitadamo *et al.* (2009), and Smith (1977) found that peer interactions in the classroom had a positive impact on the critical thinking of UGs. Conversely, Snyder and Wiles (2015) found that serving as a peer mentor had no significant impact on critical thinking when mentors were compared with nonmentors. Beyond these types of studies, there exists a void in the literature regarding whether UGs who teach/mentor younger audiences (i.e., not peer or near-peer learning) enhance their own critical thinking skills. The few studies that are available on UG mentors typically focus on UGs who are pre-service teachers and typically do not measure changes in critical thinking (Malone *et al.*, 2002).

Study Rationale

While some studies examine the effect of serving as a mentor from the UGs' perspectives, the gap in the literature becomes especially pronounced upon review of the methods used in published studies, which consist primarily of qualitative, self-reported data being used to determine the impact on the UGs in various mentoring or teaching/teaching assistant roles (Holmes *et al.*, 2013; Tenenbaum *et al.*, 2014; Ward *et al.*, 2014; Everhard, 2015; Walsh *et al.*, 2015). While self-reported data are valuable as a good starting point for research or in-depth qualitative understanding of a phenomenon, they can be considered unreliable or biased and are listed as a limitation in many studies (Linn *et al.*, 2015; Owen, 2017). Furthermore, qualitative data may not permit researchers to fully gauge how mentoring impacts specific skills such as critical thinking, which can be difficult to measure empirically (Gellin, 2003). Moreover, the length of time that an individual participates in an intervention (i.e., one semester vs. 1 year or more) may also play a role, as Snyder and Wiles (2015) did not find a significant gain in the critical thinking of peer mentors after they served as mentors for one semester, but there remains a gap in the literature involving interventions (inclusive of teaching by UGs) lasting longer time periods.

In the rare case that quantitative data are present in a published mentoring study, they typically are not the result of use of comprehensively tested instruments (Hannafin *et al.*, 1997). This suggests that there is abundant opportunity for quantitative data collection and analysis in the mentoring literature, particularly studies that employ valid and reliable instruments. A recent study by Dowd *et al.* (2018) serves as a model example of such a quantitative study, as it employs the CCTST to examine the impacts of scientific writing on the critical thinking of UGs; however, the focus of this study is based on a classroom intervention and not on mentoring youth.

Taken together, the studies to date investigate how we should teach students (at all levels) to think critically, and how peer interactions may impact the critical thinking of UGs, but there exists a void in the literature regarding whether UGs who teach/mentor younger audiences enhance their critical thinking skills. Studies that use valid and reliable instruments for this measurement are noticeably absent. Therefore, we suggest that this is the first report using a well-validated assessment (the CCTST) of a link between teaching experiences by UGs who mentor younger audiences and critical thinking skills.

RESEARCH QUESTION AND STUDY DESIGN

In this study, we focused on the UG mentoring and teaching component of NE STEM 4U. This study aims to understand whether UG mentors demonstrated gains in critical thinking after at least two semesters of mentoring to K–8 students when compared with nonmentor UGs, using the CCTST. All of the individuals within both groups (mentors and nonmentors) were life science majors at the University of Nebraska at Omaha (UNO) who took similar courses during their matriculation and were comparable in distribution of gender, ethnicity, and class standing (Supplemental Table 1). The nonmentor life science UGs served as a control group and took the CCTST at the same time periods as the mentors. Using these two groups, this study was informed by the following research questions:

1. Does serving as a mentor impact the overall critical thinking of UG mentors compared with nonmentor life science UGs, as indicated by pre/post CCTST scores?
2. Are there specific subscales of the CCTST that indicate significant differences between mentor and nonmentor life science UGs?

METHODS

The NE STEM 4U program began in 2013, and since then, UGs have learned about NE STEM 4U at new student orientation, through flyers in hallways, or on university-sponsored student-group pages. The mentees come from participating K–8 schools that have a 1-hour after-school "enrichment" time, during which the UGs deliver an inquiry-based lesson in the form of experiments. UGs pick up participating students at the school cafeteria and take them to a classroom to carry out experiments and lessons. The UGs are solely in charge of the youth during the 1-hour enrichment window and serve as a replacement of public school staff during this time (as such cannot exceed a ratio of 15 students to 1 UG). Additional programmatic information, lesson plans, and other materials to replicate the program can be found at https://nebraskaomaha.orgsync.com/org/nestem4u.

The demographics of the mentor participants in NE STEM 4U are ~85% UGs and 15% graduate students; however, UGs are the focus of this study. UG mentors have an incoming GPA range of 1.5 to 4.0 (on a 4.0 scale) and most have declared a major in a STEM discipline. More detailed information about the general mentor characteristics is found in Nelson *et al.* (2017). Detailed demographic information regarding the participants in the current study is available in Supplemental Table 1. Additionally, the demographics of the youth (i.e., K–8 participants) are included in Cutucache *et al.* (2016) and Leas *et al.* (2017), but are summarized here: >50% of participants are African American, Latino, or declare both African American and Latino as their race, and just under 50% are Caucasian, Asian, or Pacific Islander students. All schools served in the Omaha Public School District have free or reduced lunch rates >47%, with the majority of the schools served being at 97% free and reduced lunch rates or higher (Leas *et al.*, 2017).

Experimental Design

This quasi-experimental pre/posttest study used quantitative data from the CCTST to test the hypothesis that mentoring positively influenced the critical thinking of mentors in the NE STEM 4U program at the UNO when compared with nonmentor life science UGs. Groups were normalized to ensure matches for year in school, prior course work, and completion of science course work ($n = 37$). Informed consent was collected from all voluntary participants in accord with institutional review board regulations (IRB# 548-12-EX). This study took place over two academic years, with the same groups (NE STEM 4U mentor life science majors and nonmentor life science majors, respectively) and phases (quantitative pre/post) but different UGs each year.

Both mentor and nonmentor life science UGs took the CCTST at the beginning and end of the academic year (i.e., after two semesters of mentoring and course work or two semesters of course work only, respectively). The CCTST is a roughly 50-minute electronic assessment that provides an overall critical thinking score in addition to eight subscale scores: analysis, interpretation, inference, evaluation, explanation, induction, deduction, and (an optional test) numeracy. See Table 1 for a detailed definition of each measure provided by Insight Assessment (2017). The test is consistently updated based upon input from experts in fields such as assessment, psychometrics, measurement, statistics, and decision sciences, among others, and is based on the recommendations of the Delphi Report (Insight Assessment, 2017). According to the test designers, the subscores are not intended to represent completely independent factors; however, because many of the subscores are not inherently discrete units, they work together to represent the overall critical thinking ability of the student (Insight Assessment, 2017).

The questions used in the CCTST to measure reasoning skills come from a question pool that has been tested for over two decades by international measurement experts (Insight Assessment, 2017). This test is unique, because it is the only instrument that measures both cognitive and metacognitive skills, as recommended in the Delphi Report (Facione, 1990), and has been extensively evaluated for validity and reliability. A commonly cited definition of validity was provided by Eisenhart and Howe (1992, p. 1) as "the trustworthiness of inferences drawn from data." In other words, how well does an instrument measure what it is thought to measure? Reliability is generally defined as "the degree to which an assessment tool produces stable and consistent results" (American Educational Research Association *et al.*, 1985, p. 11).

Notably, many sources report on the robust validity of the CCTST (Williams *et al.*, 2003; Sorensen and Yankech, 2008; O'Hare and McGuinness, 2015). Reliability tests for the eight subscales resulted in Cronbach's alpha values ranging from 0.71 to 0.80 and a Cronbach's alpha of >0.9 for the overall instrument (Facione and Facione, 1997), scores that indicate a strong instrument (Miller and Salkind, 2002). Additionally, the test

TABLE 1. CCTST scores (overall plus eight subscales) used for this study[a]

Score	Description
Overall	How well do students use reason to inform judgment?
Analysis	Students identify how arguments are formed based on assumptions, reasons, and claims. Students also glean information from tables, figures, and documents.
Interpretation	Students resolve the precise meaning and significance of text or tables and figures; may involve clarifying, categorizing or determining significance.
Inference	Students draw probable conclusions based on reason and evidence.
Evaluation	Students determine the credibility of sources and claims.
Explanation	Students describe/articulate evidence, reasons, methods, rationales, and conclusions.
Induction	Students draw inferences about what is likely true as a basis for action.
Deduction	Students make precise, rigorously logical decisions based on specific contexts.
Numeracy	Students interpret figures and tables that present data quantitatively. They make judgments based on analysis and evaluation of mathematical/statistical information.

[a]Summarized from Insight Assessment, 2017.

has been used internationally across a wide variety of fields, including education research, science, nursing, psychology, and engineering, among others (Insight Assessment, 2017).

Analytical Procedures

All statistical tests were completed using SAS v. 9.4 (SAS Institute, Cary, NC). Before data collection, we estimated the sample size required to detect an effect using a power level of 80% and statistical significance cutoff of $p \leq 0.05$ for this study ($n = 11$ mentors; $n = 26$ nonmentors; total n value = 37). After data were collected, we tested them for normality using the Anderson-Darling test. Scores of the subscale analysis were transformed using a reciprocal transformation to achieve normality. To assess whether mentors and nonmentors differed in the pretest, we conducted one-way analyses of variance (ANOVAs) with the effect group (i.e., mentor vs. nonmentor). To investigate whether mentors improved more than nonmentors, we conducted repeated-measures ANOVAs, which included the effects test (pre vs. post; the repeated measure), group (mentor vs. nonmentor), and interaction between test and group. Specifically, we used the interaction between test and group to test for differences in improvement between mentors and nonmentors.

RESULTS

At the beginning of the academic year (Supplemental Table 2), the overall CCTST score on the pretest did not significantly differ between life science students who were NE STEM 4U mentors or nonmentors, $F(1, 35) = 3.32$, $p = 0.0771$. However, mentors scored higher in the subscales inference, $F(1, 35) = 4.92$, $p = 0.0332$; interpretation, $F(1, 35) = 5.18$, $p = 0.0291$; and numeracy, $F(1, 35) = 4.51$, $p = 0.0409$ (Supplemental Table 2).

The repeated-measures ANOVAs showed that mentors increased their scores substantially in the subscales analysis, inference, and numeracy, while nonmentors showed no change in their scores (Table 2 and Figure 1; raw data in Supplemental Table 2). Although not significant, the overall score in the CCTST test, as well as in all the other subscales, showed a visually similar pattern on average (Figure 1, Supplemental Table 2, and Supplemental Figure 1), with mentors showing an increase in all scales from pre- to posttest and nonmentors demonstrating little to no change, or negative change, from pre- to posttest (Supplemental Table 2 and Supplemental Figure 1). Test (i.e., pre- vs. posttest) and group (i.e., mentor vs. nonmentor) were significant within the repeated-measures ANOVA for the overall score and some subscores (Table 2 and Figure 1). However, this is largely because of the pull of the mean; therefore, we present the ANOVAs to ensure a focus on the specific, significant gains.

DISCUSSION

The overarching objective of this study was to determine whether participation in the NE STEM 4U intervention (i.e., the professional development program for UG STEM majors) led to significantly improved gains in critical thinking skills. Specifically, we had two research questions: 1) Does serving as a mentor impact the overall critical thinking skills of mentors (compared with nonmentors)? 2) Are there specific subscales of the CCTST that indicate significant differences between mentors and nonmentors? For the first question, serving as a mentor does not statistically significantly impact student gains in critical thinking. However, mentoring does lead to statistically significant gains for student participants in analysis, inference, and numeracy—three subscales of critical thinking. We found that, in terms of overall critical thinking score, serving as a mentor did not have a significant impact, although a marginal increase was observed. Moreover, mentoring did statistically significantly impact the critical thinking subscale scores of analysis, inference, and numeracy.

Interestingly, previous studies (Madison, 2002; Golbeck *et al.*, 2005) and the summaries of the skills (listed in Table 1) indicate a degree of relatedness between these subscale measures. Specifically, the three subscales of analysis, inference, and numeracy all relate to mathematical skills or quantitative literacy (Madison, 2002). Abilities in analysis and inference are also considered to indicate a higher level of quantitative literacy than basic numeracy or basic computational ability (Golbeck *et al.*, 2005).

TABLE 2. Descriptive statistics for NE STEM 4U mentor and nonmentor life science majors who participated in this study

	N	Mean overall pretest score ± SE	Mean overall posttest score ± SE
NE STEM 4U Mentors	11	78.55 ± 2.87	82.27 ± 1.76
Nonmentors	26	73.19 ± 1.52	73.73 ± 1.51

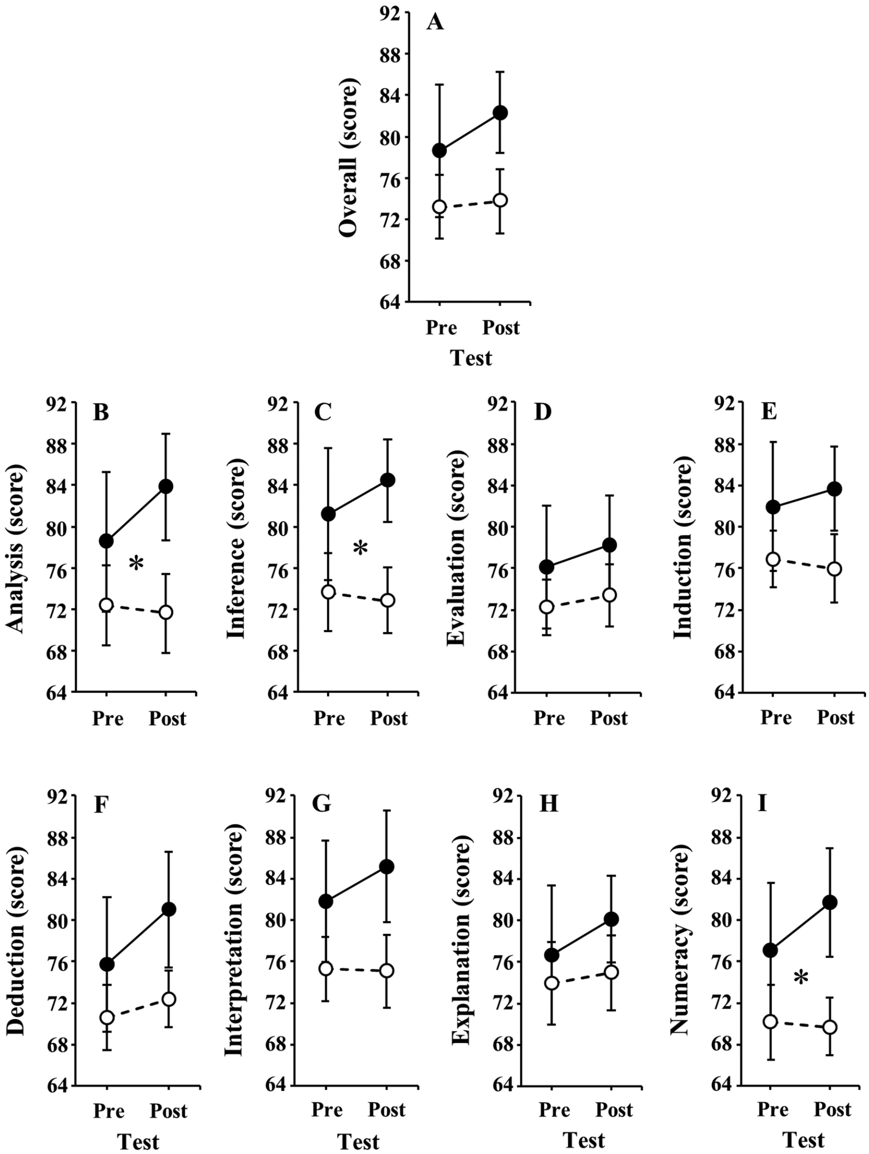

FIGURE 1. Results of repeated-measures ANOVA comparing change in performance between pre- and posttest of NE STEM 4U mentors (closed circles) to nonmentors (open circles) for overall scores (A) and eight subscales: analysis (B), inference (C), evaluation (D), induction (E), deduction (F), interpretation (G), explanation (H), and numeracy (I). Means and 95% confidence intervals are shown. An asterisk (*) indicates a significant interaction ($p \leq 0.05$) between test (pre- vs. posttest) and group (mentor vs. nonmentor). Between pre- and posttest, mentors increased their scores substantially in the subscales analysis, inference, and numeracy, while nonmentors showed no change in their score.

While it is not completely clear why UGs who mentor K–8 youth would show significant gains in measures related to math, the fact that mentors did display these gains post-mentoring is important, as studies indicate math skills are a strong predictor of future success (Trapmann *et al.*, 2007). Trapmann *et al.* (2007) found that math grades were good predictors of future success for math, engineering, and natural science majors. Notably, Trapmann *et al.* (2007) found that, for engineering students, math grades were better predictors of academic success than an aptitude test specific to engineering. While the current study involved life science majors and not engineering students, it is interesting to note that mentoring significantly improved critical thinking abilities related to math skills.

In the current study, the observed increases in mathematical skills and quantitative literacy are likely due to the structure of the lesson plans/curricula for the NE STEM 4U program, as these are approximately equally balanced throughout the academic year to include math lessons in addition to science lessons. Moreover, many science lessons also include heavy use of mathematics, so this may be a by-product of frequency of exposure to and practice with these principles. This should be further explored in future studies. Additionally, more work should be done to understand why mentors did not demonstrate significant gains in the subscale scores of interpretation, evaluation, explanation, induction, and deduction. Specifically, the subscale of explanation, at least intuitively, seems to be an area that would be heavily used by a mentor/teacher of younger audiences, yet mentors did not show significant gains in this area. These questions remain to be discerned in future studies with larger sample sizes.

Another significant question on what is driving the improvement in critical thinking scores is whether it is the process of teaching itself (“learning better by teaching”) or whether it is the fact that the participants in the NE STEM 4U program are in a group dynamic (“camaraderie encouraging improvement”). The latter has been demonstrated by Springer *et al.* (1999) in a meta-analysis of decades of data on STEM UGs. We recognize the challenges in trying to tease apart the contribution of the process of teaching on the UGs as compared with the group dynamic. We have several studies ongoing with different cohorts of students (some in cohesive groups, others not) who serve as teachers for youth, and we expect to be able to address this limitation in future work, but it remains a significant barrier for understanding the precise contribution (if such a phenomenon can be determined) herein.

Overall, the findings in this study provide evidence that mentoring in NE STEM 4U improved critical thinking of the mentors when compared with nonmentor life science UGs, but more work needs to be done to further understand and corroborate these findings. For example, the findings of this study would be more robust if we had: a larger sample size, additional mentoring programs outside of NE STEM 4U, and a broader variety of STEM majors from different universities included. Additionally, the length of time that UGs participate in similar interventions (i.e., one semester vs. 1 year or more) should be further investigated to determine whether mentoring duration plays a role in critical thinking development, as Snyder and Wiles (2015) did not find a significant gain in the critical thinking of peer mentors after they served as mentors for one semester.

However, these preliminary findings do strongly suggest that serving as a UG mentor can improve critical thinking. Therefore, encouraging UGs to serve as mentors may be a way to fulfill the 21st-century skill development that many researchers say courses and other experiences are not meeting (Singer *et al.*, 2012; NACE, 2014). In addition, serving as a UG mentor significantly improved quantitative skills such as analysis, inference, and numeracy, which are known to be strong indicators of future success for UGs in academics and their future careers (Trapmann *et al.*, 2007). Overall, this quantitative study supports the findings of a previous qualitative study, wherein former UG mentors self-reported that they felt their experience improved their critical thinking (Nelson and Cutucache, 2017). More studies such as these should be conducted to provide strong empirical evidence of the impact serving as a mentor has on UG mentors.

We suggest that the incorporation of an innovative model that provides transferable skills to UGs for future employment, coupled with gains in critical thinking skills to apply to their course work and then ultimately to meet community stakeholder needs is a win-win-win. Finally, the levels of retention to academic degree completion as well as placement in the STEM workforce for NE STEM 4U mentors were significantly higher (i.e., 95.9%, as reported in the *Introduction*) than the national average, thus suggesting the importance of this program for recruitment and retention of STEM majors. Overall, this study suggests that serving as a teacher/mentor to younger audiences may lead to gains in specific subscore or components of critical thinking for UGs.

ACKNOWLEDGMENTS

We thank the Sherwood Foundation and the Peter Kiewit Foundation for support of this project. Additionally, thanks to all student participants in the project and community partners such as Collective for Youth, Beyond School Bells, and Omaha Public Schools.

REFERENCES

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, DC.

Association of American Colleges and Universities. (2013). It Takes More Than a Major: Employer Priorities for College Learning and Student Success. Retrieved July 1, 2018, from www.aacu.org

Bailey, R., & Mentz, E. (2015). IT teachers' experience of teaching-learning strategies to promote critical thinking. *Issues in Informing Science and Information Technology*, *12*, 141–152.

Carneval, A. P., Smith, N., & Stoll, J. (2010). Help wanted: Projections of jobs and education requirements through 2018. Retrieved July 1, 2018, from https://cew.georgetown.edu/jobs2018

Carpenter, S. (2015). Undergraduates' perceived gains and ideas about teaching and learning science from participating in science education outreach programs. *Journal of Higher Education Outreach and Engagement*, *19*(3), 113–146.

Cutucache, C. E., Luhr, J., Nelson, K., Grandgenett, N., & Tapprich, W. (2016). NE STEM 4U: An out-of-school time academic program to improve achievement of underprivileged youth in STEM areas. *International Journal of STEM Education*, *3*(6), 1–7.

Dowd, J. E., Thompson, R. J., Jr., Schiff, L. A., & Reynolds, J. A. (2018). Understanding the complex relationship between critical thinking and science reasoning among undergraduate thesis writers. *CBE—Life Sciences Education*, *17*, ar4.

Eisenhart, M., & Howe, K. (1992). Validity in educational research. In LeCompte, M., Millroy, W., & Preissle, J. (Eds.), *The handbook of qualitative research in education*. San Diego: Academic.

Everhard, C. J. (2015). Implementing a student peer-mentoring programme for self-access language learning. *Studies in Self-Access Learning Journal*, *6*(3), 300–312.

Facione, N. C., & Facione, P. A. (1997). *Critical thinking assessment in nursing education programs: An aggregate data analysis. A research report*. Millbrae, CA: California Academic Press.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. Millbrae, CA: California Academic Press.

Gellin, A. (2003). The effect of undergraduate student involvement on critical thinking: A meta-analysis of the literature 1991–2000. *Journal of College Student Development*, *44*(6), 746–762.

Golbeck, A. L., Ahlers-Schmidt, C. R., Paschal, A. M., & Dismuke, S. E. (2005). A definition and operational framework for health numeracy. *American Journal of Preventive Medicine*, *29*, 375–376.

Greenwald, R. R., & Quitadamo, I. J. (2014). A mind of their own: Using inquiry-based teaching to build critical thinking skills and intellectual engagement in an undergraduate neuroanatomy course. *Journal of Undergraduate Neuroscience Education*, *12*(2), A100–A106.

Gutek, G. L. (2009). *New perspectives on philosophy and education*. Hoboken, NJ: Pearson.

Hannafin, M. J., Hannafin, K. M., Land, S. M., & Oliver, K. (1997). Grounded practice and the design of constructivist learning environments. *Educational Technology Research and Development*, *45*(3), 101–117.

Holmes, N. G., Martinuk, M., Ives, J., & Warren, M. (2013). Teaching assistant professional development by and for TAs. *Physics Teacher*, *51*, 218–219.

Holyoak, K. J., & Morrison, R. G. (2005). *The Cambridge handbook of thinking and reasoning*. New York: Cambridge University Press.

Insight Assessment. (2017). *CCTST users manual and guide*. San Jose: California Academic Press.

Langdon, D., McKittrick, G., Beede, D., Khan, B., & Doms, M. (2011, July 1). STEM: Good Jobs Now and for the Future (ESA Issue Brief #03-11). Washington, DC: U.S. Department of Commerce, Economics and Statistics Administration.

Leas, H., Nelson, K., Grandgenett, N., Tapprich, W., & Cutucache, C. (2017). Fostering curiosity, inquiry, and scientific thinking in elementary school students: impact of the NE STEM 4U intervention. *Journal of Youth Development*, *12*(2), 103–120. doi: https://doi.org/10.5195/jyd.2017.474

Linn, M. C., Palmer, E., Baranger, A., Gerard, E., & Stone, E. (2015). Undergraduate research experiences: Impacts and opportunities. *Science*, *347*(6222), 1261757.

Madison, B. L. (2002). Educating for numeracy: A challenging responsibility. *Notices of the American Mathematical Society*, 181.

Magrabi, S. A. R., Pasha, M. R., & Pasha, M. Y. (2018). Classroom teaching to enhance critical thinking and problem-solving skills for developing IOT applications. *Journal of Engineering Education Transformations*, *31*(3), 152–157.

Malone, D., Jones, B.D., & Stallings, D. T. (2002). Perspective transformation: Effects of a service-learning tutoring experience on prospective teachers. *Teacher Education Quarterly*, *29*(1), 61–81.

Miller, D., & Salkind, N. (2002). *Handbook of research design and social measurement* (6th ed.). London: Sage.

National Academies of Sciences, Engineering, and Medicine. (2016). *Developing a national STEM workforce strategy: A workshop summary*. Washington, DC: National Academies Press. doi: 10.17226/21900

National Association of Colleges and Employers (NACE). (2014). Job outlook 2015. Bethlehem, PA. Retrieved July 17, 2015, from www.umuc.edu/upload/NACE-Job-Outlook-2015.pdf

National Science Board. (2010). *Preparing the next generation of STEM innovators: Identifying and developing our nation's human capital*. Arlington, VA: National Science Foundation.

Nelson, K. L., & Cutucache, C. E. (2017). How do former undergraduate mentors evaluate their mentoring experience 3-years post-mentoring: A phenomenological study. *Qualitative Report*, *22*(7), 2033–2047.

Nelson, K. L., Sabel, J., Forbes, C., Grandgenett, N., Tapprich, W., & Cutucache, C. (2017). How do undergraduate STEM mentors reflect upon their mentoring experiences in an outreach program engaging K-8 youth? *International Journal of STEM Education*, *4*(3), 1–13.

New York Academy of Sciences. (2014). The global STEM paradox. A report of the New York Academy of Sciences, 2014. Retrieved July 1, 2018, from https://globalstemalliance.org/media/filer_public/b8/68/b8683358-027a-4015-8807-811301744bbd/nyas_white_papers_ssf.pdf.

O'Hare, L., & McGuinness, C. (2015). The validity of critical thinking tests for predicting degree performance: A longitudinal study. *International Journal of Educational Research*, *72*, 162–172.

Owen, L. (2017). Student perceptions of relevance in a research methods course. *Journal of Applied Research in Higher Education*, *9*(3), 394–406.

Paris, B. (2016). Failing to improve critical thinking. *Inside Higher Ed*. Retrieved April 17, 2018, from www.insidehighered.com/views/2016/11/29/roadblocks-better-critical-thinking-skills-are-embedded-college-experience-essay

Phillips, V. & Bond, C. (2004). Undergraduates' experiences of critical thinking. *Higher Education, Research and Development, 23*(3), 277–294.

Quitadamo, I. J., Jayne Brahler, C., & Crouch, G. J. (2009). Peer-led team learning: A prospective method for increasing critical thinking in undergraduate science courses. *Physical Therapy Faculty Publications*, *18*(1), 29–39.

Rowe, M. P., Gillespie, B. M., Harris, K. R., Koether, S. D., Shannon, L-J. Y., & Rose, L. A. (2015). Redesigning a general education science course to promote critical thinking. *CBE—Life Science Education*, *14*(3), ar30.

Singer, S. R., & Nielsen, N. R. & Schweingruber, H. A., (Eds.), (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. Washington, DC: National Academies Press.

Smith, D. G. (1977). College classroom interactions and critical thinking. *Journal of Educational Psychology*, *69*(2), 180–190.

Snyder, J. J., & Wiles, J. R. (2015). Peer led team learning in introductory biology: Effect on peer leader critical thinking skills. *PLoS One*, *10*(1), e0115084.

Sorensen, H. A., & Yankech, L. R. (2008). Precepting in the fast lane: Improving critical thinking in new graduate nurses. *Journal of Continuing Education in Nursing*, *39*(5), 208–216.

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research*, *69* (1), 21–51.

Tenenbaum, L. S., Anderson, M. K., Jett, M., & Yourick, D. L. (2014). An innovative near-peer mentoring model for undergraduate and secondary students: STEM focus. *Innovative Higher Education*, *39*, 375–385.

Trapmann, S., Hell, B., Hirn, J. W., & Schuler, H. (2007). A meta-analysis of the relationship between the big five and academic success at university. *Zeitschrift fur Psychologie*, *215*(2), 132–151.

Walsh, D. S., Veri, M. J., & Willard, J. J. (2015). Kinesiology career club: Undergraduate student mentors' perspectives on a physical activity-based teaching personal and social responsibility program. *Physical Educator*, *72*, 317–339.

Ward, G. E., Thomas, E. E., & Disch, W. B. (2014). Mentor service themes emergent in a holistic undergraduate peer-mentoring experience. *Journal of College Student Development*, *55*(6), 563–580.

Watanabe-Crockett, L. (2018, April 4). 12 strong strategies for effectively teaching critical thinking skills. *Wabisabi* [blog]. Global Digital Citizen Foundation. Retrieved April 17, 2018, from https://globaldigitalcitizen.org/12-strategies-teaching-critical-thinking-skills

Williams, K. B., Glasnapp, D. R., Tilliss, T. S., Osborn, J., Wilkins, K., Mitchell, S., … Schmidt, C. (2003). Predictive validity of critical thinking skills for initial clinical dental hygiene performance. *Journal of Dental Education*, *67*(11), 1180–1192.