Assessing Mainstream Pre-service Teachers’ Self-Efficacy to Teach English Language Learners

Yao Fu
Asst. Prof., corresponding author, University of Wisconsin-Whitewater, USA, fuy@uwu.edu

Jiaxi Wang
University of Wisconsin-Stevens Point, USA, jwang@uwsp.edu

This study intended to provide evidence of the reliability and validity of the ELL Education Self-Efficacy Scale. Embedded in Bandura’s self-efficacy theory, the scale was created to assess mainstream pre-service teachers’ self-efficacy in teaching English Language Learners (ELLs). The participants of this study were pre-service teachers whose primary training has been in one or more traditional subject areas, such as mathematics, science, English, or social studies. The content validity of the ELL Education Self-Efficacy Scale was developed and established through five steps: defining the construct, creating the domain and item pool based on the literature, deciding the format and wording, conducting expert reviews, and implementing a pilot study. Exploratory factor analysis, parallel analysis, and Cronbach’s alpha were also employed to examine data collected from 278 mainstream pre-service teachers in six American universities. The results supported a 46-item scale with three underlying factors, including sociocultural efficacy, linguistic efficacy, and pedagogical-content efficacy. The study provides educators with insights on improving teacher education to prepare all teachers to work with ELLs.

Keywords: teaching self-efficacy, ELL education, scale measurement, teachers, teaching

INTRODUCTION

With globalization and continuous trend of immigration, there has been a growing number of culturally and linguistically diverse students in many English-speaking countries, such as New Zealand, Australia, Canada, and the U.S. (Geva & Wiener, 2015). Between 1994 and 2014, the number of English Language Learners (ELLs) in American public schools, who were non-native English speakers with limited English skills, increased from about 3.2 million to 4.6 million, making up 9.4 percent of the overall student population (National Centre for Education Statistics, 2017). Despite this increasing diversity in student populations, ELLs have been marginalized in the U.S. educational system and often experience “triple segregation” by race, poverty, and
language proficiency (Heilig & Holme, 2013; Iddings et al., 2012). To provide equitable educational opportunities to all children, nurturing future teachers’ efficacy in working with diverse student groups is essential.

In the U.S. context, there has been an urgent need to study mainstream pre-service teachers’ efficacy beliefs in teaching ELLs in response to the changing student demographics and under-preparation of mainstream teachers to work with this student group (Villegas, 2018). However, researchers face two major challenges in studying ELL teaching self-efficacy: the dearth of research and limited scale instruments. First, a scarcity of research exists regarding mainstream pre-service teachers’ efficacy in working with ELLs. The lack of research on this subject can be largely explained by insufficient emphasis on preparing teachers in teaching this student group. Instead of perceiving ELL education as a shared responsibility between language specialists and mainstream teachers, many teachers believe that bilingual teachers and/or teachers certified in English for Speakers of Other Languages (ESOL) are solely responsible for teaching ELLs although ELLs spend most of their time in content-area classrooms (Yoon, 2008). To ensure that future teachers are competent to work with all students, researchers have highlighted the importance of reconceptualizing the knowledge and skill base for mainstream teachers (Bunch, 2013; Villegas et al., 2018).

Second, few sound scale instruments exist to measure mainstream pre-service teachers’ efficacy in teaching ELLs. Among those existing scales, the Culturally Responsive Teaching Self-Efficacy Scale (CRTSE; Siwatu, 2007) demonstrates limited structural validity; and the Teaching English Language Learners Scale (TELLS; Carney, 2012) was flawed by mixing up pre- and in-service teachers as participants and failed to fully validate the two-factor structure through confirmatory factor analysis. Given that pre- and in-service teachers displayed significant differences in their beliefs about responsibility for ELLs’ learning as well as their readiness and self-competence to work with ELLs in content-area classes (Polat, 2010; Polat & Mahalingappa, 2013), the two groups should be treated as distinctive populations. This demands a scale that specifically measures mainstream pre-service teachers’ efficacy in teaching ELLs.

Considering that limited research has studied the ELL education self-efficacy in content-area classrooms (Villegas et al., 2018), we attempted to bridge this gap in the literature by developing a new scale. Different from Carney’s (2012) approach to treat pre- and in-service teachers as the same population, the current study aimed to develop a scale dedicated to assessing mainstream pre-service teachers. Even though the CRTSE (Siwatu, 2007) was created for the preservice teacher population, the scale failed to demonstrate sufficient structural validity because all those multiple-factor solutions generated from the statistical analysis were not interpretable. Thus, the current study intended to explore a factor solution which can be meaningfully interpreted and shows improved structural validity. Additionally, the CRTSE and TELLS do not adequately address on the linguistic self-efficacy which demands pre-service teachers to develop foundational linguistic knowledge to work with ELLs, such as understanding basic language components, second language acquisition, and differences between academic and conversational language (Lucas et al. 2008; Tran, 2014). By contrast, the current
study has added the linguistic domain in the scale development based on a systematic review of the relevant literature, which provides a measurement tool to better capture the construct. In detail, the article begins with a description of literature that serves as a foundation for this study. It then outlines rigorous steps to develop and establish the content validity of the ELL Education Self-Efficacy Scale; and examines the structure validity and internal reliability of the scale through analyzing survey data of 278 pre-service teachers in six American universities. The study concludes with results and implications for enhancing teacher education to prepare all teachers in working with diverse students.

Literature Review

This section provides a brief overview of the literature that established the theoretical underpinning of this study. It describes teacher self-efficacy, literature pertinent to scale measurements of teacher efficacy and domains of ELL education teacher efficacy, which guided the design of the scale.

Construct of Teacher Efficacy

Two research strands dominate the study of teacher efficacy: In the early studies, teacher efficacy was associated with teachers’ internal attributions of their teaching effectiveness, which was built on Rotter’s (1954) locus of control theory; in more recent research, Bandura’s (1997) self-efficacy theory has been widely adopted as the theoretical lens to capture the construct of teacher efficacy. In this study, we employed Bandura’s definition of self-efficacy to provide a clear theoretical foundation. Bandura (1997) defined self-efficacy as “beliefs in one’s capabilities to organize and execute the courses of action required to produce given attainments” (p. 3). Grounded in his definition, teacher self-efficacy refers to teachers’ beliefs in their own competence to achieve certain teaching tasks. Given that self-efficacy, in nature, is task- and context-specific (Bandura, 1997), teachers’ perceived efficacy may be varied based on the subjects which they are teaching, the group of students they are working with, and the settings where teaching and learning transpire. The context-specific nature of the self-efficacy demands researchers to take those factors into consideration while developing teacher self-efficacy scales (Bandura, 2006). Despite the existence of two strands on teacher efficacy, it has been supported that teachers’ efficacy beliefs can considerably impact variables related to teaching and learning. For instance, researchers found that high levels of teacher efficacy could contribute to students’ academic, affective, and motivational gains (Capara et al., 2006; Love et al., 2020; Maguire, 2011). Teacher efficacy also strongly correlates with teachers’ sense of burnout, degree of job satisfaction, and effectiveness of instruction (Huangfu, 2012; Kasalak & Dagyar, 2020; Skaalvik & Skaalvik, 2010).

Scale Measurements of Teacher Self-Efficacy

As the effect of teacher efficacy is being widely recognized, various measurements have been created to assess the construct. Three types of scales dominate the measurements of teacher self-efficacy, including scales of teacher self-efficacy in general, in teaching certain subject area(s), and in working with diverse students. The following section

International Journal of Instruction, July 2021 ● Vol.14, No.3
reviews eight representative scales which adopted Bandura’s self-efficacy theory and followed the psychometric scale development procedure. Among the generic scales of teaching self-efficacy, two well-established instruments are the Ohio State Teacher Efficacy Scale (OSTES; Tschannen-Moran & Hoy, 2001) and the Teachers’ Efficacy Beliefs System-Self Form (TEBS-Self; Dellinger et al., 2008). The OSTES was flawed by including both pre- and in-service teachers as participants, which failed to use consistent samples to develop the scale. In comparison, the TEBS-Self was questionable in structure validity in that three different factor-solutions were generated. However, both instruments include items that measure teacher efficacy in pedagogical skills, such as classroom management.

Building upon studies on teacher self-efficacy in general, more researchers realized that scale development regarding teacher self-efficacy should be more context-specific. That is, generic measurements of teacher self-efficacy cannot thoroughly investigate how teachers display differing efficacy beliefs in teaching different subject areas and working with diverse students. As a result, the specificity level of recent scales has significantly increased. For instance, the Self-Efficacy for Teaching Mathematics Instrument (SETMI; McGee & Wang, 2014) is geared towards measuring elementary teachers’ self-efficacy in teaching mathematics. Grounded in the SETMI, Alkaharusi and his co-authors (2017) adjusted the content to align with the math curriculum in the Sultanate of Oman and developed a scale to assess teachers’ self-efficacy in teaching 5-10th grade math. Apart from measuring math teaching self-efficacy, instruments have been developed in other content areas. For example, Kocabas, Ozfidan, and Burlbaw (2018) developed a measurement of teachers’ self-efficacy in teaching compulsory K-12 theology courses while Yoon, Evans, and Strobel (2014) created the Teaching Engineering Self-Efficacy Scale to assess K-12 teachers’ self-efficacy beliefs involved in engineering education. Despite measuring different subject areas, those scales underscore the importance of teacher self-efficacy in teaching content knowledge in addition to pedagogical skills.

Different from scales to measure subject-area teaching efficacy, the Culturally Responsive Teaching Self-Efficacy Scale (CRTSE; Siwatu, 2007) appraises pre-service teachers’ self-efficacy beliefs in teaching diverse students; and it includes items to assess teachers’ sociocultural efficacy in recognizing the discontinuity between students’ home and school cultures and working towards reducing the cultural mismatch. Unfortunately, none of the multiple-factor solution based on eigenvalues and scree test were interpretable. By contrast, the Teaching English Language Learners Scale (TELLS; Carney, 2012) contains two factors: “support and resources” and “assessment and instruction”, which were not fully supported by confirmatory factor analysis. Both the CRTSE and TELLS consist of a few items that touch upon teachers’ linguistic self-efficacy in using students’ native language to greet and praise ELLs; however, the linguistic domain is not adequately addressed due to insufficient emphasis on linguistic efficacy, such as teaching language irregularities, distinguishing between academic and conversational language proficiency, and applying knowledge related to second language acquisition to support ELL learning. In brief, those scales provided guidance to generate items and factor structure of the ELL Education Self-Efficacy Scale.
It is clear from the literature that research focusing on teacher self-efficacy has attracted more attention in the last two decades. Despite the large number of scales to assess teaching self-efficacy in general and content-area classrooms, limited studies have probed mainstream pre-service teachers’ self-efficacy in teaching ELLs. Given that pre-service teachers can display various self-efficacy beliefs while working with different student populations (Siwatu, 2011), it is important to create a scale to more accurately capture their efficacy beliefs related to ELL education in content-area classrooms.

**Domains of ELL Education Self-Efficacy**

Given that self-efficacy is highly context-specific, we also examined relevant literature on ELL education to ensure that the development of scales “must be tailored to the particular domain of functioning that is the object of interest” (Bandura, 2006, p. 308). We reviewed the Education Commission’s (2014) comparison of 50 states’ teaching standards. An analysis of standards from states that require ELL training for teachers (e.g., Alabama, New Jersey, New York, and Washington) reveals three primary knowledge sources of teaching ELLs, including pedagogical-content, linguistic, and sociocultural domains. Specifically, mainstream pre-service teachers should develop pedagogical-content knowledge to make content accessible to ELLs, linguistic knowledge to facilitate language acquisition, and sociocultural knowledge to promote ELL cultural identity. Additionally, the literature review of textbooks and journal articles on ELL education confirms and elaborates those three primary knowledge sources as unfolded below.

**Pedagogical-content Domain.** Pedagogical-content knowledge is essential for teacher efficacy in working with ELLs because it not only requires mastery of various teaching methods and adequate content knowledge, but also demands teachers to meaningfully integrate them into their daily teaching practices to support all learners (Shulman, 1987). Researchers have supported that effective assessment, instructional strategies, classroom management, and rigorous curriculum may play a pivotal role in teaching diverse students. In terms of assessment, researchers, like Gottlieb (2006), highlighted effective measurement of ELLs’ learning outcomes must include an evaluation of two parts: language proficiency and academic achievement. That is, teachers should differentiate between students’ limited English skills and learning difficulties in content areas and assess those two aspects respectively.

Apart from effective assessments, teachers should be competent in using a variety of instructional strategies. Hill and Flynn (2006) recommended six types of classroom instruction that works with this student group, which include (1) setting clear learning objectives and giving timely constructive feedback to ELLs, (2) using non-linguistic activities to help ELLs formulate and elaborate on knowledge, (3) asking questions that are focused on important aspects of learning and allowing students enough time to react to the questions, (4) selecting appropriate grouping strategies, and encouraging cooperative learning, (5) enhancing students’ summarizing and note-taking skills, and (6) assigning meaningful homework and practice for students to apply and reinforce their acquired knowledge.
Building a conducive classroom learning environment is another dimension in which teachers should develop strong efficacy. Celic (2009) highlighted the need to establish consistent daily schedules so that students can better understand their teachers’ expectations and more easily adjust to the learning environment. Based on students’ performance, teachers need to modify the classroom routines to meet the special needs of ELLs. To ensure productive learning, teachers should also set up clear classroom rules that “govern how students act in the classrooms” (Celic, 2009, p. 78). Once the rules are collaboratively determined by the teacher and students, teachers must address the importance of following these rules and explain how students will be held accountable for their behaviors in class.

Lastly, teachers should be efficacious in providing ELLs with “a challenging and creative curriculum which enables students to acquire disciplinary knowledge and promote higher-order thinking skills” (Garcia & Kleifgen, 2010, p. 73). Such a curriculum can help ELLs understand the key ideas and concepts of the content area, apply acquired knowledge and skills to a new context, build connections between concrete know-how with abstract theories, and develop critical thinking skills to digest knowledge and information (Gibbons, 2009). To successfully create and implement the curriculum, teachers have to integrate bilingual learning materials or materials in students’ primary language, use culturally relevant books, and tap various technological resources appropriate for ELLs (Freeman, 2007; Quirocho & Ulanoff, 2009).

Linguistic Domain. With the increasing language variety displayed in schools, mainstream teachers need to be linguistically competent to effectively work with ELLs. Instead of requiring teachers to gain mastery in linguistics or speak a second language, all classroom practitioners should have foundational linguistic knowledge, including knowledge on second language acquisition (Bunch, 2013). Fillmore and Snow (2000) described this linguistic competence as “educational linguistics” that “would support teachers’ undertakings overall, and in particular teaching literacy skills and working with English language learners” (p. 5). They outlined some core aspects of teachers’ linguistic knowledge, including basic language components, language irregularities, word formation, technical vocabulary, dialect variation, academic language, second language acquisition, English orthography, writing systems, and rhetorical structures of different writings. Similarly, Lucas et al. (2008) advocated that cultivating linguistic competence is crucial for teacher education; and they concluded that linguistically responsive teachers should be able to (1) distinguish between conversational language proficiency and academic language proficiency, (2) ensure ELLs to understand the information conveyed to them and offer opportunities to practice their language skills, (3) support language growth of ELLs through active interaction with others, (4) encourage ELLs to use their native language, (5) create a classroom environment conducive to the learning of all students, and (6) provide explicit instruction on linguistic knowledge necessary to second language learning.

Sociocultural Domain. Beyond pedagogical-content and linguistic competence, teachers should possess capabilities to work closely with ELL students, their families, and communities to address the social and cultural needs of this student group. Apart

International Journal of Instruction, July 2021 ● Vol.14, No.3
from the pressure of developing English proficiency and academic-content knowledge simultaneously, ELLs are stressed to adapt to the school culture which may be quite different from their own cultural background. From a sociocultural theoretical perspective, the educational experience of ELLs is not only concerned with cognitive development, but also a social process in which these students reconcile differences in the minority and dominant cultures, and gradually form their sociocultural identity (Alsagoff, 2012; Reyes & Ervin-Tripp, 2010). In particular, August and Shanahan (2006) summarized that “being a member of a low-status language group may have negative effects on self-concept, motivation and/or learning opportunities, all of which can depress literacy attainment” (p. 258). This finding resonates Boone’s (2011) study, which identifies one of the important reasons for ELLs dropping out of school as “not feeling part of the social circle” (p. 432). In congruence with these study results, Yoon (2008) illustrated how ELLs were disengaged and alienated in mainstream classrooms without teachers’ support and accommodations. To better attend to the socioemotional development of ELLs, teachers should build strong family-school connections to boost students’ self-confidence, facilitate their adaption to new cultures, and create supportive learning environments in class and at home (Amatea, 2012; Geva & Wiener, 2015).

**The Current Study.** We intended to develop a scale to assess mainstream pre-service teachers’ self-efficacy in working with ELLs and examine the validity and reliability of the scale. The study population included pre-service teachers who planned to teach in P-12 content-area classrooms in their future careers. Embedded in Bandura’s (1997) self-efficacy theory, the term teacher self-efficacy refers to teachers’ beliefs in their own capabilities in achieving given tasks. The three sources of teacher knowledge described by existing literature provide a sound basis for developing the scale and conceptualizing the key domains of teacher self-efficacy in ELL instruction. In detail, the Pedagogical-content Domain refers to pre-service teachers’ abilities in interweaving various teaching methods and content knowledge to support ELLs learning in subject areas; the Linguistic Domain describes how pre-service teachers take full use of their fundamental linguistic knowledge to improve ELLs’ English proficiency; and the Sociocultural Domain is defined as pre-service teachers’ capabilities to effectively work with ELLs, their families, and communities to identify and accommodate to the sociocultural needs of diverse students.

**METHOD**

In the current study, we initially developed a draft of the ELL Education Self-Efficacy Scale with 52 items to assess mainstream pre-service teachers’ beliefs in their own capacity to teach diverse learners. To establish the content validity, we defined the construct, created the domain and item pool based on the literature, decided the format and wording, conducted expert reviews, and implemented a pilot study with ten mainstream pre-service teachers to refine the draft scale (DeVellis, 2012). Then, we recruited 278 mainstream pre-service teachers in six American universities to complete the revised version of the ELL Education Self-Efficacy Scale; and employed exploratory factor analysis, parallel analysis, and Cronbach’s alpha to uncover the underlying factor structure and examine the internal reliability of the scale (Furr & Bacharach, 2014;
O’Connor, 2000; Worthington & Whittaker, 2006). The following section details the development of the draft scale, the demographic information of the participants, the study procedure, and data analysis.

Instrument Development

Following DeVellis’ (2012) guidelines, we undertook five rigorous steps to develop and establish the content validity of the scale. First, the scale was established on Bandura’s (1997) self-efficacy theory, which provided a solid theoretical underpinning to conceptualize what was being measured. Given that self-efficacy beliefs can vary significantly in different contexts, we intended to specifically investigate mainstream pre-service teachers’ self-efficacy in working with ELLs.

Second, we identified the scale domains and created an item pool based on a careful examination of the relevant literature. As described in the literature review section, three major types of information were reviewed for domain and item generations: (1) six representative psychometric scales on teacher self-efficacy, (2) journal articles and textbooks on ELL education, and (3) the teaching license requirements from states which specifically demand teachers to receive training in teaching ELLs (Education Commission, 2014). By synthesizing the various sources of information, we strived to create scale items that are grounded in research and teaching practices and can accurately measure the construct.

Third, we adopted the 100-point scale format and appropriate wording of the items following Bandura’s (2006) recommendations. According to Pajares et al. (2001), the 100-point scale enables researchers to better measure respondents’ differentiating judgements, which is considered as a more sensitive and reliable scale format than the five-point Likert-scales. In addition, we phrased the items in terms of can do, because “can is a judgment of capability; will is a statement of intention” (Bandura, 2006, p. 43).

Fourth, we invited fifteen experts to help review the scale and improve the content validity, which included faculty in higher education, in-service teachers and school administrators with expertise in ELL education. They not only went through the survey, but also offered insights on scale revision. The panel of experts contributed to revising the directions and improving the wording and content of individual items. Several items were modified to clarify the meaning. For instance, item one was originally written as “I can use appropriate assessments to evaluate the English skills of ELLs”. The experts unanimously reported that it would be more appropriate to revise it as “I can differentiate assessments to evaluate the English skills and academic learning of ELL students”. The reason of such a change was that mainstream pre-service teachers, in most cases, assessed students’ learning outcomes in content areas, and the differentiation between assessments of language and academic skills is more essential in content-area classroom settings.

Fifth, we conducted a pilot study to pretest the scale. Ten pre-service teachers were recruited from a medium-sized university to complete the scale and discuss the relevance, clarity, and conciseness of the items. Then, we refined the scale based on preservice teachers’ feedback, such as clarifying the meaning of non-linguistic activities.
by adding a couple of examples of such activities in the parenthesis to make the content more transparent to the audience. The refined scale with 52 items was used as the instrument of the study.

Participants
We recruited 278 participants from six public universities located in the northeast regions of the U.S. They had to meet two criteria: (1) They were at least 18 years old and (2) they would not be licensed in ESOL and/or bilingual education. The average age of the pre-service teachers was twenty-one years old, with the youngest being eighteen and the oldest forty years of age. About 85% of the participants were female, and 15% were male. The vast majority of the participants were Caucasians with less than 6% of the pre-service teachers having a different racial background. Eighty-seven percent of the participants self-reported as monolingual English speakers, and 13% were proficient in a second language. The participants were from different years in college, 22.7% freshman, 18.3% sophomore, 41.7% junior, 15.1% senior, and 2.2% other. Approximately 64.4% of the participants were early childhood education majors, 25.5% elementary/middle education majors, and 10.1% secondary education majors.

Instruments & Procedure
A background questionnaire was firstly employed to collect demographic information of the 278 participants, including age, gender, ethnicity, language proficiency, and year in the current program. Then, the participants responded to the draft of the ELL Education Self-Efficacy Scale (Appendix A) which consisted of 52 items. The definition of teaching ELL self-efficacy was presented at the beginning of the survey, which provided clear concepts to the participants. The body of the survey was constructed to explore teachers’ self-efficacy in alignment with the three domains of ELL instruction as described in the literature review: the pedagogical content domain, linguistic domain, and sociocultural domain. An example of scale items under the pedagogical content domain was “I can use non-linguistic activities to help ELL students formulate and elaborate on knowledge”. In terms of the linguistic domain, the participants responded to questions, such as “I can use my knowledge in second language acquisition to support ELL learning”. For the sociocultural domain, a sample item was “I can understand the challenges and anxieties that ELLs may undergo in adapting to a different culture”.

Data Analysis
The survey data from 278 participants were analysed by the SPSS software to explore the underlying factor structure. Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) and Bartlett’s Test of Sphericity were conducted to ensure that the sample was adequate and appropriate for factor analysis. Then, we applied exploratory factor analysis (EFA) and parallel analysis to determine the retention or extraction of the factors. To improve interpretation, Direct Oblimin rotation was selected to reduce cross-loadings given that the three factors were correlated to each other (Zygmont & Smith, 2014). Cronbach’s alphas of items within each factor were also computed to examine the internal consistency of the scale (Furr & Bacharach, 2014). The content validity of the scale was established by taking rigorous steps to develop the instrument, including
defining the construct, creating the domains and item pool based on the literature, conducting expert reviews, and piloting the scale with pre-service teachers. The EFA and parallel analysis results supported the ELL education self-efficacy as a three-factor construct and showed promise of the structural validity of the scale. The high Cronbach’s alpha within all three factors indicated excellent internal reliability of the scale.

FINDINGS

The analysis results supported a three-dimensional scale with 46 items. The KMO value was .96, indicating the sample size was sufficient in performing EFA. The Barlett’s Test of Sphericity $\chi^2 (1326) = 16043.97, p < 0.01$ was significant, which suggested the correlations between variables were different from zero. Specifically, we used factor analysis to examine the underlying latent dimensions of the self-efficacy construct. Given that the dataset’s normality (i.e., Kurtosis, Skewness, Kolmogorov-Smirnov and Shapiro-Wilk) was violated, principal factor analysis was applied as the estimation method because it requires no distributional assumptions and is less likely than maximum likelihood to generate a solution with a Heywood case (Fabrigar, Wegener, MacCallum, & Strahan, 1999). We also conducted a parallel analysis, a more robust method suggested by O’Connor (2000) to extract the number of factors. The result of the parallel analysis showed that the eigenvalue of the fourth factor in the raw data (i.e., 1.61 scale mean) is smaller than its counterpart (i.e., 1.79 scale mean at 95 percentile) associated with the parallel analysis, and thus three factors were retained. Deletion of individual items from the factor structure relied on four criteria: (1) items with all factor loadings less than .40, (2) items with cross-loadings less than .15 difference from an item’s highest factor loading, (3) items with absolute loadings higher than .40 on two or more factors, and (4) items with factor loading that cannot be meaningfully interpreted (e.g., the meaning of the item is not related to the factor on which it is loaded) (Worthington & Whittaker, 2006). Guided by these criteria, 6 items were deleted, which led to a scale with 46 items, which constitute the final version of the ELL Education Self-Efficacy Scale. Table 1 presents detailed information on the rotated factor loadings of the 46 items retained in the scale.
Table 1
Rotated factor loadings of the forty-six items retained in the scale

<table>
<thead>
<tr>
<th>Item No.</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.865</td>
<td>-.017</td>
<td>-.104</td>
</tr>
<tr>
<td>2</td>
<td>.930</td>
<td>-.075</td>
<td>-.085</td>
</tr>
<tr>
<td>3</td>
<td>.941</td>
<td>.093</td>
<td>-.017</td>
</tr>
<tr>
<td>4</td>
<td>.849</td>
<td>.083</td>
<td>-.103</td>
</tr>
<tr>
<td>5</td>
<td>.843</td>
<td>.083</td>
<td>-.040</td>
</tr>
<tr>
<td>6</td>
<td>.873</td>
<td>-.003</td>
<td>.056</td>
</tr>
<tr>
<td>7</td>
<td>.853</td>
<td>.065</td>
<td>-.002</td>
</tr>
<tr>
<td>8</td>
<td>.853</td>
<td>-.077</td>
<td>.118</td>
</tr>
<tr>
<td>9</td>
<td>.596</td>
<td>.033</td>
<td>.217</td>
</tr>
<tr>
<td>10</td>
<td>.654</td>
<td>.235</td>
<td>.075</td>
</tr>
<tr>
<td>11</td>
<td>.681</td>
<td>-.116</td>
<td>.114</td>
</tr>
<tr>
<td>12</td>
<td>.623</td>
<td>.267</td>
<td>.053</td>
</tr>
<tr>
<td>13</td>
<td>.557</td>
<td>.113</td>
<td>.310</td>
</tr>
<tr>
<td>14</td>
<td>.565</td>
<td>.226</td>
<td>.173</td>
</tr>
<tr>
<td>15</td>
<td>.522</td>
<td>.324</td>
<td>.104</td>
</tr>
<tr>
<td>16</td>
<td>.542</td>
<td>.372</td>
<td>-.015</td>
</tr>
<tr>
<td>17</td>
<td>.532</td>
<td>.366</td>
<td>.066</td>
</tr>
<tr>
<td>19</td>
<td>.488</td>
<td>.271</td>
<td>.199</td>
</tr>
<tr>
<td>22</td>
<td>.120</td>
<td>.665</td>
<td>.127</td>
</tr>
<tr>
<td>23</td>
<td>.250</td>
<td>.540</td>
<td>.144</td>
</tr>
<tr>
<td>24</td>
<td>.082</td>
<td>.657</td>
<td>.096</td>
</tr>
<tr>
<td>25</td>
<td>-.061</td>
<td>.824</td>
<td>-.017</td>
</tr>
<tr>
<td>26</td>
<td>.023</td>
<td>.717</td>
<td>-.173</td>
</tr>
<tr>
<td>27</td>
<td>-.022</td>
<td>.623</td>
<td>.217</td>
</tr>
<tr>
<td>28</td>
<td>.133</td>
<td>.724</td>
<td>.084</td>
</tr>
<tr>
<td>30</td>
<td>.081</td>
<td>.745</td>
<td>.087</td>
</tr>
<tr>
<td>31</td>
<td>.130</td>
<td>.749</td>
<td>.022</td>
</tr>
<tr>
<td>32</td>
<td>.025</td>
<td>.760</td>
<td>.125</td>
</tr>
<tr>
<td>33</td>
<td>.135</td>
<td>.716</td>
<td>.065</td>
</tr>
<tr>
<td>34</td>
<td>.102</td>
<td>.705</td>
<td>-.010</td>
</tr>
<tr>
<td>37</td>
<td>.096</td>
<td>.010</td>
<td>.752</td>
</tr>
<tr>
<td>38</td>
<td>.114</td>
<td>.034</td>
<td>.693</td>
</tr>
<tr>
<td>39</td>
<td>.130</td>
<td>-.040</td>
<td>.784</td>
</tr>
<tr>
<td>40</td>
<td>.109</td>
<td>-.227</td>
<td>.867</td>
</tr>
<tr>
<td>41</td>
<td>.077</td>
<td>-.243</td>
<td>.867</td>
</tr>
<tr>
<td>42</td>
<td>.014</td>
<td>.118</td>
<td>.743</td>
</tr>
<tr>
<td>43</td>
<td>-.041</td>
<td>.007</td>
<td>.847</td>
</tr>
<tr>
<td>44</td>
<td>-.041</td>
<td>.044</td>
<td>.883</td>
</tr>
<tr>
<td>45</td>
<td>-.062</td>
<td>-.009</td>
<td>.893</td>
</tr>
<tr>
<td>46</td>
<td>-.087</td>
<td>.008</td>
<td>.931</td>
</tr>
<tr>
<td>47</td>
<td>-.149</td>
<td>.078</td>
<td>.866</td>
</tr>
<tr>
<td>48</td>
<td>.046</td>
<td>.126</td>
<td>.625</td>
</tr>
<tr>
<td>49</td>
<td>.125</td>
<td>.055</td>
<td>.639</td>
</tr>
<tr>
<td>50</td>
<td>-.005</td>
<td>.137</td>
<td>.727</td>
</tr>
<tr>
<td>51</td>
<td>-.024</td>
<td>.168</td>
<td>.739</td>
</tr>
<tr>
<td>52</td>
<td>-.045</td>
<td>.175</td>
<td>.657</td>
</tr>
</tbody>
</table>

Note: Items 18, 20, 21, 29, 35 & 36 were deleted from the draft of the ELL Education Self-Efficacy scale.
As revealed from the EFA analysis, this three-factor scale with 46 items explained about 67.78% of the total variance in the data. Consistent with the scale conceptualization, the three factors were named as the Pedagogical-Content Self-Efficacy, Linguistic Self-Efficacy, and Sociocultural Self-Efficacy respectively. To begin with, the first factor was Pedagogical-Content Self-Efficacy, which included eighteen items. Some items loaded in this factor included “I can use a variety of assessments to track ELL students’ academic achievement in content areas”, and “I can use non-linguistic activities (e.g. visual and kinaesthetic activities) to help ELL students formulate and elaborate on knowledge”. Two items (i.e., items 18 and 20) were deleted because they had cross-loadings less than .15 difference from an item’s highest factor loading. The Cronbach’s alpha was .96, indicating excellent internal reliability within factor one.

Linguistic Self-Efficacy was the second factor revealed from the analysis, which consisted of 12 items. Some examples of the retained items were “I can motivate ELL students to pay attention to language irregularities that may confuse them”, and “I can understand the English variability displayed by ELL students, such as vernacular dialects and accents”. Four items (i.e., items 21, 29, 35 and 36) were eliminated due to high cross-loadings on two factors. The internal reliability of the eleven items under factor two was also high, α = .95.

The last factor was the Sociocultural Self-Efficacy. All the sixteen items originally created within this factor were kept on the scale. These items included “I can understand the challenges and anxieties that ELL students may undergo in adapting to a different culture”. Another example was “I can build connections between ELL students’ cultural background and their classroom learning experiences”. The Cronbach’s alpha was .96, demonstrating the highest internal consistency among the three factors. In summary, the substantial amount of variance accounted by the 46 items along with the excellent internal consistency within each factor supported the three-factor structure of the scale.

DISCUSSION

With the increasingly diverse student populations in the U.S., the P-12 education should be strengthened to promote the social and cognitive development of all learners, such as ELLs (Council for the Accreditation of Education Preparation, 2020; Takanishi & Menestrel, 2017). This study sheds light on how to revitalize ELL education in mainstream classrooms. To begin with, the development of the ELL Education Self-Efficacy Scale is a step forward in understanding the underlying dimensions of the construct. The validity of test content was established by defining the construct to be measured, creating domains and items based on systematic literature review, conducting expert reviews, and piloting the scale with a sample of pre-service teachers. The internal reliability was examined using Cronbach’s alpha. Results from EFA and parallel analysis showed promise of structural validity, supporting the three domains of ELL education identified from the literature. Despite the large number of measures of self-efficacy, few sound scale instruments exist to assess mainstream pre-service teachers’ self-efficacy in educating ELLs. Further, researchers (Huangfu, 2012; Kasalak & Dagyar, 2020; Love et al., 2020; Mojavezi & Tamiz, 2012; Skaalvik & Skaalvik, 2010; Viel-Ruma, Houchins, Jolivette, & Benson, 2010) found that teachers’ self-efficacy
beliefs have a significant impact on actual teaching performance as well as student academic achievements and learning motivation. Given the need of more sound instruments in ELL education and a strong influence of self-efficacy on teaching and learning, the three-factor scale uncovered by the current study, even in its nascent form, opens new possibilities for research. Specifically, the three domains establish a foundation to further examine the reliability and validity of the current scale: The first domain is pre-service teachers’ self-efficacy in synthesizing content knowledge and pedagogical skills to maximize ELL students’ academic achievement in content areas; the second domain is to nurture their efficacy in developing fundamental linguistic knowledge and fully use such knowledge to boost students’ English proficiency; the third domain is enhancing their efficacy beliefs in identifying the sociocultural needs of ELLs, and collaborating with families and communities to bridge the home and school cultures. Compared to the TELLS (Carney, 2012) and CRTSE (Siwatu, 2007), the current scale has more comprehensively examined the construct of ELL education self-efficacy by including the linguistic domain, an essential area that has been overlooked or underplayed in existing measurements.

In addition, the results offer suggestions to enhance the quality of teacher education programs. In the U.S., only 14 states have requirements related to ELL education; and among those states, there has been no generally accepted approach or teaching standards regarding how to prepare pre-service teachers in teaching ELLs in content-area classrooms (Education Commission, 2014; Hughes & Mahalingappa, 2018). As a result, the wide variation in different teacher preparation programs has created ambiguity in learning and instruction. In most cases, the topics related to ELL learning were randomly embedded in several courses without in-depth discussions on practical teaching methods effective for ELL students; and thus, many pre-service teachers’ learning experience on ELL instruction has been fragmentary and superficial (Durgunoglu & Hugh, 2010; Jimenez-Silva, Olson, & Hernandez, 2012). In more recent studies, specially designed interventions have been created and embedded in a semester-long course to enhance mainstream pre-service teachers’ beliefs and efficacy related to ELL teaching (Polat et al., 2019; Yough, 2019); however, the content and approach of such interventions differ substantially and the long-term effectiveness of those interventions have not been probed. By reviewing studies on the preparation of future mainstream teachers to teach ELLs, Villegas and her co-authors (2018) have not only confirmed the limited studies on this research area, but also highlighted “the absence of studies that focused on preparing future teachers to analyse the language demands embedded in academic text and learning tasks, an indispensable skill to scaffold instruction adequately for ELLs” (p. 152). To optimize the quality of existing programs, teacher educators should be aware of the interactions among the pedagogical content, linguistic, and sociocultural domains. In addition, the individual items of the scale provide thoughts on some essential concepts and skills instructors may address in class to effectively enhance all pre-service teachers’ efficacy in teaching diverse students. While representing different categories of teachers’ knowledge base, the three domains should be meaningfully integrated to achieve effective ELL teaching. This idea also echoes Lee’s (2004) advocate of instructional congruence which refers to “the process
of merging academic disciplines with students’ linguistic and cultural experiences to make the academic content accessible, meaningful, and relevant for all students” (p. 66). The integration of the three domains challenge mainstream pre-service teachers to strengthen their self-efficacy in teaching content-area knowledge, facilitating ELLs to develop language proficiency, and attending to their sociocultural needs.

Lastly, the current scale might serve as the basis to develop new scales that measure in-service teachers’ self-efficacy of ELL teaching in content-area classrooms. Because of inadequate preparation in teacher education programs, it was not surprising to find many in-service teachers were not ready to teach diverse student populations (Pettit, 2011). Some in-service teachers were reluctant to work with ELLs and believed that teaching these students was the sole responsibility of English as a second/foreign language or bilingual teachers; some demonstrated biased and limited understanding about ELLs and their families; others, willing to work with this student population, were not confident about their teaching competence (Walker, Shafer, & Liams, 2004; Yoon, 2008). For instance, Pappamihiel and Lynn (2016) found that many mainstream in-service teachers could not clearly distinguish between instructional and linguistic accommodations for ELLs while acknowledging the importance of knowing the difference. The dichotomy between teachers’ belief and competence demands a thorough investigation of the linguistic self-efficacy among in-service teachers. Apart from their limited linguistic efficacy, Cruz and her co-authors (2020) reported that in-service teachers were less confident in their cultural knowledge than building trust and personal relationships with diverse students. Given that in-service teachers are not adequately prepared to teach ELLs, it is essential to use the ELL Education Self-Efficacy Scale to more accurately measure teachers’ pedagogical-content self-efficacy, linguistic self-efficacy, and sociocultural self-efficacy and fully capture elements in each domain that need to be strengthened. In specific, the current scale could be used as an informal self-assessment tool to help teachers reflect on their teaching practice through analysis of their own strengths and weaknesses. Additionally, school administrators (e.g., principals and superintendents) can easily apply or adapt the current scale to collect a large number of in-service teachers’ responses about their self-efficacy beliefs in teaching ELLs. This first-hand information directly garnered from teachers could inform instructional and curricular improvement and identify needs in professional development. In turn, this large sample of data collected from the mainstream in-service teachers can be used to further test the reliability and validity of the current scale with that population.

**CONCLUSION**

All students, regardless of their cultural backgrounds, need an equitable environment where they could achieve their best potential. This holds true for the increasing number of ELLs who are often marginalized in American content-area classrooms. Despite an abundance of literature on measuring teaching self-efficacy in general and in subject areas, there is a meagre of sound scale instruments that specifically assess mainstream pre-service teachers’ self-efficacy in working with ELLs. We attempt to bridge this gap by developing the ELL Education Self-Efficacy Scale. The analysis of 278 pre-service teachers’ data supports a 46-item scale with three underlying factors evidenced by
content and structure validity and internal reliability. The findings of this study shed light on enhancing teacher education and professional development to prepare all pre- and in-service teachers to work with ELLs.

LIMITATIONS
The current study could be improved in two aspects. The participants of the study were recruited from universities in the northeast region of the U.S., so the sample might not sufficiently represent the overall pre-service teacher population across the country. Thus, the generalizability of the study might be limited by geographic location, sample size, and participants’ educational experience in teacher education programs. Given this limitation, researchers can use a larger and more representative sample by recruiting pre-service teachers across the U.S. In addition, future research could further examine the validity and reliability (e.g., convergent and divergent validity) of the ELL Education Self-Efficacy Scale given that the current study focused on initial scale development and validation. Study efforts could also be delineated towards uncovering the applicability of the scale with in-service teachers as well as comparing and contrasting findings between pre- and in-service teacher populations.

REFERENCES


**Appendix A. Draft of the ELL Education Self-Efficacy Scale**

**Directions:** This scale includes 52 items that ask about how you perceive your own capabilities in working with English language learners (ELLs). In this study, **ELL students** refer to non-native English speakers who are not yet proficient in English and require instructional support in order to thrive in regular classrooms. As a pre-service teacher who will teach in mainstream classrooms, please rate your competence in ELL instruction on a 100-point scale with 0 representing “Cannot do at all” and 100 representing “Highly certain can do”. You should answer these questions to the best of your knowledge by circling the choices that most accurately reflect your current situation. All your responses will be anonymous.

1. I can differentiate assessments to evaluate the English skills and academic learning of ELL students.
2. I can use a variety of assessments to track ELL students’ academic achievement in content areas.
3. I can accommodate ELL students’ learning needs while planning and administering assessments.
4. I can skilfully analyse and interpret ELL students’ assessment results.
5. I can provide constructive feedback to ELL students based on their assessment results.
6. I can use appropriate grouping strategies to engage ELL students in collaborative learning.
7. I can set clear learning goals for ELL students.
8. I can use non-linguistic activities (e.g. visual and kinaesthetic activities) to help ELL students formulate and elaborate on knowledge.
9. I can provide enough wait time for ELL students to respond to my questions.
10. I can assign meaningful homework and practice for ELL students to apply and reinforce knowledge taught in class.
11. I can establish consistent schedules to help ELL students adjust to the classroom environment.
12. I can use a variety of strategies to manage ELL students’ disruptive behaviors and other special needs/difficulties.
13. I can help ELL students practice the classroom routines through modeling.
14. I can create clear classroom rules and communicate them effectively to ELL students.
15. I can effectively work with ELL students to address their discipline problems.
16. I can develop higher-order thinking skills in my ELL students through teaching the curriculum.
17. I can provide ELL students with a curriculum that is challenging and creative.
18. I can offer ELL students rich learning resources that are bilingual or written in their native language.
19. I can integrate abundant technological resources in the curriculum to help ELL students learn more effectively.
20. I can use learning materials that are culturally relevant to ELL students’ life.
21. I can understand different units of language, such as morpheme, phoneme, words and phrases.
22. I can motivate ELL students to pay attention to language irregularities that may confuse them.
23. I can teach some basic principles of word formation to aid ELL students’ vocabulary acquisition.
24. I can understand the English variability displayed by ELL students, such as vernacular dialects and accents.
25. I can use my metalinguistic knowledge to analyze the similarities and differences between English and other languages.
26. I can use my knowledge in second language acquisition to support ELL students’ learning.
27. I can create opportunities for ELL students to speak in their native language.
28. I can ensure that ELL students understand the information conveyed to them in class.
29. I can allow enough time for ELL students to practice different English skills, including listening, speaking, reading and writing.
30. I can distinguish between ELL students’ academic language proficiency and social/communicative language proficiency.
31. I can explicitly teach academic terminologies that are challenging to ELL students.
32. I can explicitly teach text structures to ELL students, such as differences in narratives and expository texts.
33. I can address the discourse patterns and rhetorical devices pertinent to academic tasks.
34. I can accumulate some knowledge in ELL students’ native language, such as greetings.
35. I can develop realistic expectations for ELL students based on their current language skills in English and their native language.
36. I can show empathy and support for ELL students who experience hardships.
38. I can understand the challenges and anxieties that ELL students may undergo in adapting to a different culture.
39. I can boost the self-confidence and self-esteem of ELL students.
40. I can appreciate the cultures and values that ELL students bring to the class.
41. I can promote diversity and mutual respect in class.
42. I can maximize opportunities for ELL students interacting with their English-proficient peers.
43. I can help my English-speaking students deepen their understanding about other cultures.
44. I can spend time on helping ELL students develop a sense of belonging to the learning community.
45. I can encourage ELL students to make their voice heard.
46. I can accumulate knowledge of ELL students’ home culture.
47. I can develop a critical and impartial perspective of ELL students’ families.
48. I can maintain frequent communications with ELL students’ parents.
49. I can work collaboratively with parents to address the special needs of ELL students.
50. I can recognize the discontinuity between ELL students’ home culture and school culture.
51. I can build connections between ELL students’ cultural background and their classroom learning experiences.
52. I can identify rich resources in the local communities to support ELL students’ learning.