Measurement in Physical Education and Exercise Science, 16: 284–299, 2012

ISSN: 1091-367X print / 1532-7841 online
DOI: 10.1080/1091367X.2012.716726

Routledge
Taylor & Francis Group

Development of a Physical Education Teaching Efficacy Scale

Charlotte A. Humphries and Edward Hebert
Department of Kinesiology and Health Studies, Southeastern Louisiana University, Hammond, Louisiana

Kay Daigle
Department of Health, Physical Education, and Recreation, Southeastern Oklahoma State University, Durant, Oklahoma

Jeffrey Martin
Division of Kinesiology, Health and Sport Studies, Wayne State University, Detroit, Michigan

Relationships have been found between teacher efficacy and many teaching and learning variables, but few researchers have examined teaching efficacy in physical education. The instrument reported here, the Physical Education Teaching Efficacy Scale, was developed based on the teaching efficacy literature, existing scales, and National Association for Sport and Physical Education's Teacher Education Standards. Students attending 11 institutions who are majoring in regular and alternate physical education teacher education and who are at different stages of preparation (N = 592) completed the initial survey. Exploratory, followed by confirmatory, factor analysis resulted in a 35-item, 7-factor scale. Factors were Content Knowledge, which were activities one might teach; Applying Scientific Knowledge in Teaching, which reflected academic content; Accommodating Skill Differences; Teaching Students with Special Needs; Instruction, which included management, motivation, and instruction; Using Technology; and Assessment. Results support that the Physical Education Teaching Efficacy Scale addresses many aspects of teaching physical education and meets research criteria for validity and reliability.

Key words: physical education, efficacy, teacher education, instrument development

INTRODUCTION

Self-efficacy beliefs reflect one's capabilities to exercise control over events and estimations of competence to execute given tasks. Efficacy beliefs affect performance, influence the selection

Correspondence should be sent to Charlotte A. Humphries, Department of Kinesiology and Health Studies, Box 10845, Southeastern Louisiana University, Hammond, LA, 70402. E-mail: Charlotte.Humphries@selu.edu

of tasks, and are a key factor in self-regulation of motivation. Over the past 30 years, scholars have studied efficacy beliefs of teachers and reported relationships between a teacher's sense of efficacy and a multitude of desirable outcomes, including student achievement (see Tschannen-Moran, Hoy, & Hoy, 1998, for a review). In physical education (PE) research, teacher self-efficacy has been linked to professional development (Martin, McCaughtry, Hodges-Kulinna, & Cothran, 2008; Martin, McCaughtry, Kulinna, Cothran, & Faust, 2008), and teacher and student behavior (Martin & Hodges-Kulinna, 2004; Martin & Kulinna, 2005). As a result of findings such as these, teaching efficacy has been identified as a powerful factor in the teaching and learning process (Ashton, 1984; Guskey & Passaro, 1994; Soodak & Podell, 1996; Tschannen-Moran et al., 1998).

Research on teacher efficacy can be traced to two Rand Corporation assessments of educational programs in which researchers found that teachers' efficacy beliefs were positively related to student achievement and to the likelihood that teachers would follow through on implementing new strategies (Armor et al., 1976; Berman, McLaughlin, Bass-Golod, Pauly, & Zellman, 1977). These studies used two Likert-scaled items to assess teacher efficacy: "When it comes right down to it, a teacher really can't do much, because most of a student's motivation and performance depends on his or her home environment;" and "If I try really hard, I can get through to even the most difficult or unmotivated students." Later, Ashton and her colleagues (Ashton, 1984; Ashton, Webb, & Doda, 1983) used multiple data sources (the two items from the Rand studies, interviews, and classroom observations) and found a relationship between teacher efficacy and student achievement, identified differences between low- and high-efficacy teachers, and described workplace factors that undermined teaching efficacy.

During this era, Gibson and Dembo (1984) developed the Teacher Efficacy Scale, which conceptualized two aspects of the construct: "personal teaching efficacy"—a teacher's expectations of being able to bring about learning in students, and "general teaching efficacy"—one's beliefs regarding the extent to which teachers in general can overcome outside factors that impede student learning. Though there has been much discussion and debate, both conceptually and in terms of measurement, the Teacher Efficacy Scale and variations of it have been the dominant means by which teaching efficacy has been assessed (see Henson, 2002; Klassen,Tze, Betts, & Gordon, 2011; Tschannen-Moran et al., 1998 for reviews).

Researchers have consistently shown relationships between a teacher's sense of efficacy and assorted variables in the teaching and learning process. For example, teacher efficacy has been related to classroom management strategies (Ashton et al., 1983), willingness to adopt new ideas (Ghaith & Yaghi, 1997; Stein & Wang, 1988), persistence to work with students who struggle (Gibson & Dembo, 1984), expectations for student performance (Tournaki & Podell, 2005), teacher evaluations (Riggs & Enochs, 1990), and faculty absenteeism (Imants & Van Zoelen, 1995). A teacher's level of efficacy has also been shown to be positively related to student expectations for their own performance (Anderson, Greene, & Loewen, 1988), as well as student achievement (Armor et al., 1976; Caprara, Barbaranelli, Steca, & Malone, 2006; Dembo & Gibson, 1985). Finally, teaching efficacy has been shown to have implications for job satisfaction, commitment to the profession, and length of career (Caprara et al., 2006; Glickman & Tamashiro, 1982; Trentham, Silvern, & Brogdon, 1985).

While researchers have shown that teacher efficacy has great promise, many questions remain to be answered. For example, in their review, Tschannen-Moran et al. (1998) urged researchers to examine factors that contribute to the development of teacher efficacy, study the stability of efficacy over career stages, and examine how efficacy varies with contextual factors such as student

characteristics and subject matter. The study of teacher efficacy has also been criticized for its treatment of the construct as a global disposition (Bandura, 1997; Wheatley, 2005). The primary instruments that have been used to measure teacher efficacy have failed to ground the construct in Bandura's (1986, 1997) conceptualization of efficacy as task and/or situation specific and may therefore be inadequate. In line with this idea, researchers have begun to develop teaching efficacy instruments specific to content areas, such as science (Riggs & Enochs, 1990), and writing (Graham, Harris, Fink, & MacArthur, 2001). Most recently, multi-dimensional instruments have been constructed that examine teachers' efficacy beliefs specific to certain aspects of the teaching process, such as knowing the course content, motivating students, adapting instruction to student needs, assessing students, managing behavior, and overcoming barriers to student engagement (e.g., Baker, 2005; Brouwers & Tomic, 2000; Duncan & Ricketts, 2008; Martin & Kulinna, 2003; Skaalvik & Skaalvik, 2007; Woolfolk-Hoy & Spero, 2005).

To date, however, with the exception of Martin and colleagues, few researchers have examined teaching efficacy among physical educators, or efficacy specific to teaching PE. The one existing PE teacher self-efficacy instrument is narrowly focused, examining teachers' efficacy to teach PE lessons with high levels of physical activity (Martin & Kulinna, 2003). Hence, the purpose of this study was to develop a broader, multi-dimensional teaching efficacy instrument specific to personal teaching efficacy for PE. In a society beset by sedentary lifestyles and intolerable obesity rates (Centers for Disease Control and Prevention, 2012), information that can help increase physical educators' effectiveness is increasingly important. This inference is based on the body of literature linking higher teacher efficacy with desirable teacher behaviors—in this case, getting all students to perform moderate to vigorous physical activity in a safe, supportive environment. Teacher efficacy may be of particular concern in PE in light of findings that teacher efficacy was weaker in the presence of excessive role demands, low status, lack of recognition, professional isolation, and alienation (Tschannen-Moran et al., 1998; Webb & Ashton, 1987), characteristics that are often reported in PE. Thus, the availability of a teaching efficacy instrument specific to PE can be a significant contribution.

METHODS

Instrument Development

In keeping with the multi-dimensional instruments developed by Skaalvik and Skaalvik (2007) and Woolfolk-Hoy and Spero (2005), we sought to create an instrument that assessed efficacy for various aspects of teaching PE. A multi-dimensional approach was chosen to reflect the complexity of teaching PE.

As a starting framework for identifying important aspects of teaching PE, we used the multi-factor efficacy surveys previously developed and the Initial Physical Education Teacher Education Standards developed by the National Association for Sport and Physical Education (NASPE, 2001, 2009). The NASPE Standards address expectations related to prospective teachers' knowledge base, broad areas of teaching (such as planning, management and motivation, communication, and student assessment), and professionalism. Standards are further linked to outcomes/elements illustrating more specific expectations. The 2009 revision is briefer than the 2001 standards, but is not substantively different in its conceptualization of effective teaching.

The intent was to develop items that addressed major aspects of teaching PE and would also be understandable to the practitioner. When possible, items were worded to be clear to undergraduates, while using terminology from the standards. For example, the NASPE Standards (2001, 2009) indicate preservice teachers should be able to develop and implement appropriate instructional explanations, cues, and prompts to facilitate competent motor performance. The item written to reflect this idea was, "I can use clear teaching cues that help students remember and understand how to do a skill correctly." Expectations regarding management suggest that teachers should be able to organize, allocate, and manage resources to provide active and equitable learning experiences. To parallel this idea, we wrote the item, "I can use available space and equipment to make sure students get lots of practice and do not spend time waiting in line." Table 1 lists sample items and corresponding NASPE Standards.

In some instances, we sought to provide context for the respondent. For example, the NASPE Standards indicate physical educators should be able to plan and adapt instruction for diverse students. Items constructed to address this idea required respondents to provide efficacy estimations for effectively developing activities for low- or high-skilled students, or for students with specific disabilities. These decisions were based on the idea that efficacy judgments are likely to vary with teaching certain types of content or to students with specific characteristics. Because it was impractical to create items for all possible types of activity or all possible forms of student diversity, items were developed for situations that physical educators would be likely to encounter. This was also consistent with the recommendations of Bandura (2006), Bong (2006), and Pajares, Hartley, and Valiante (2001), that self-efficacy items be specific enough to give meaningful context, but not so specific as to preclude generalizability.

After writing and revising items, a version of the instrument was provided to three outside readers, university faculty and researchers in PE teacher education in other states and regions. They evaluated each item relative to the clarity of the statement, the extent to which each item reflected the ideas in the NASPE Standards, and each item's importance for inclusion in the instrument relative to other comparable items. Changes were made based on their feedback; two items were deleted based on this review, and many were rewritten.

Next, a convenience sample of 19 undergraduate PE teacher education candidates completed a 74-item version of the survey instrument, responded to the items, evaluated their clarity, and provided feedback on the response scale. This feedback suggested editorial changes to a few items; results from the feedback also indicated that a meaningful efficacy scale could be achieved by using a 10-point scale anchored by (1) "Disagree/Cannot do," and (10) "Agree/Highly certain I can do," with a midpoint of (5) "Neutral/Moderately certain I can do."[1]

At this point, the instrument (Physical Education Teaching Efficacy Scale [PETES]) consisted of 80 efficacy items answered using the 1–10 scale, along with a section for demographic characteristics, university attended, and status in the university's PE teacher education program. University Institutional Review Board approval was given for this and all other phases of data collection.

[1]The survey initially used a 100-point scale, as recommended by Bandura (2006). However, everyone used only multiples of 10 to rate items they scored below 70, while several candidates gave items rated above 70 with such scores as 76, 94, and so forth, suggesting they were lapsing into a grading scale, despite instructions to use mid-range numbers as neutral. This suggested that, with this population, a 100-point scale could be a threat to validity, and the more traditional 10-point scale was substituted.

TABLE 1
Examples of PETES Survey Items and Corresponding NASPE Standards

Sample PETE Item	*2001 NASPE Standard/Outcome*		*2008 NASPE Standard/Element*	
I have a good grasp of exercise science concepts (from Exercise Physiology, Biomechanics, Motor Learning, and Sport Psychology) and can apply them to teaching PE.	1. Content Knowledge	Describe and apply biophysical (anatomical, physiological, and biomechanical) and social-psychological concepts to skillful movement, physical activity, and fitness.	1. Scientific and Theoretical Knowledge	Describe and apply physiological and biomechanical concepts related to skillful movement, physical activity, and fitness.
I know how first graders are different from fourth graders physically, cognitively, socially, and emotionally.	2. Growth and Development	Understand the biological, psychological, sociological, experiential, and environmental factors (e.g., neurological development, physique, gender, socioeconomic status) that influence developmental readiness to learn and refine movement skills.	1. Scientific and Theoretical Knowledge	Describe and apply motor development theory and principles related to skillful movement, physical activity, and fitness.
If one of my students were having trouble with a drill/activity, I know ways to change it to make it easier for them.	3. Diverse Students	Identify, select, and implement appropriate instruction that is sensitive to students' strengths/weaknesses, multiple needs, learning styles, and prior experiences.	3. Planning and Implementation	Plan and adapt instruction for diverse student needs, adding specific accommodations and/or modifications for student exceptionalities.
I can use routines so students do not waste time during class.	4. Management and Motivation	Use managerial routines that create smoothly running learning experiences. Organize, allocate, and manage resources (e.g., students, time, space, equipment, activities, teacher attention) to provide active and equitable learning experiences.	4. Instructional Delivery and Management	Utilize managerial rules, routines, and transitions to create and maintain a safe and effective learning environment.
I can talk to students who are overweight, low skilled, of different races, and to both males and females so I don't insult them or hurt their feelings	5. Communication	Communicate in ways that demonstrate sensitivity to all students (e.g., considerate of ethnic, cultural, socio-economic, ability, gender differences).	6. Professionalism	Communicate in ways the convey respect and sensitivity.

I can demonstrate and explain a skill/drill so the class understands what to do.	6. Planning and Instruction	Use effective demonstrations and explanations to link physical activity concepts to appropriate learning experiences.	4. Instructional Delivery and Management	Implement effective demonstrations, explanations, and instructional cues and prompts to link physical activity concepts to appropriate learning experiences.
I can use assessments for both grading and to help me plan.	7. Student Assessment	Interpret and use learning and performance data to make informed curricular and/or instructional decisions.	5. Impact on Student Learning	Utilize assessments and reflection to foster student learning and inform instructional decisions.
I often look for ideas about teaching PE from other sources such as teachers, the internet, professional conferences, and professional literature.	8. Reflection	Use available resources (e.g., colleagues, literature, professional organizations) to develop as a reflective professional.	6. Professionalism	Participate in activities that enhance collaboration and lead to professional growth and development.
I can integrate technology if I have access to it (such as video and sound systems) into my teaching.	9. Technology	Design, develop, and implement student learning activities that integrate information technology.	3. Planning and Instruction	Demonstrate knowledge of current technology by planning and implementing learning experiences that require students to appropriately use technology to meet lesson objectives.
I can get information or equipment or use facilities for my PE class from people who work in local fitness centers and recreation departments.	10. Collaboration	Identify and actively seek community resources to enhance physical activity opportunities.	6. Professionalism	Participate in activities that enhance collaboration and lead to professional growth and development.

Participants

PE teacher education candidates at 11 institutions across the United States were asked to complete the survey. Institutions were selected on the basis of the presence of an undergraduate PE teacher education program and toward the pursuit of a diverse set of respondents. Institutions included rural and urban settings, both public and private, and ranged from small (total enrollment of 2,700) to large (total enrollment of 42,000) universities. Of the 595 completed surveys, 592 were usable. Most respondents ($n = 430$) were traditional undergraduates, and the rest (22 of whom had been teaching without a standard license) were in an alternative certification program. Of the 592 respondents, the majority were white ($n = 434$, 73%) or African American ($n = 107$, 18%), and males outnumbered females 63% to 37%.

Data Analysis and Results

Data were randomly split into two samples. We then conducted an exploratory factor analysis using a principal components analysis with a varimax rotation. Scree plots were also examined. The results revealed a seven-factor scale, with Eigenvalues greater than 1.0, accounting for 66% of the total variance. Factor loadings ranged from .42 to .77, and 53 items were retained. Internal consistency of each factor was assessed using Cronbach's alpha; these values ranged from .77 to .94, all meeting our criterion level of .70 (Cronbach, 1951). To examine test–retest reliability, another 64 respondents completed the PETES twice over a three-day period, and the reliability of each factor score was evaluated using intraclass correlation. These correlations ranged from .63 to .88.

To confirm the factor structure obtained in the first half of the split sample, we then conducted a confirmatory factor analysis (CFA) using structural equation modeling (SEM) techniques. Maximum likelihood estimation was used with AMOS 5.0 (Arbuckle, 2003). Standard conditions were specified by having the appropriate indicators load on each of the seven latent factors, and then we defined the scale of each latent factor by fixing the factor loading of one indicator for each latent variable to one. Subscales were free to correlate.

We examined a variety of model goodness of fit (GOF) indexes to evaluate different aspects of model fit (Kline, 2005). Results of the CFA indicated an inadequate fit based on the following fit indexes that did not meet commonly used criteria (Byrne, 2006; Kline, 2005): χ^2 ($df = 1{,}253$, $N = 297$) = 3,319.28, $p < .001$, $\chi^2/df = 2.65$, Comparative Fit Index (CFI) = .80, Root Mean Square Error of Approximation (RMSEA) = .08 and 90% confidence interval = .072–.078. These findings suggested that a stronger fit might be achieved with model modification. Furthermore, to develop a more practical scale (i.e., reduce participant burden) we also sought to reduce the total number of scale items. Because factor five (i.e., efficacy about instruction) had far more items (20) than all of the other factors, we were particularly interested in deleting items from it. At the same time, because factor five had items reflecting management, motivation, and communication, we did not want to eliminate the items that might represent one particular area.

Kline (2005) suggested that both empirical (e.g., factor loadings) and substantive considerations (e.g., theory, logic) be used in the re-specification of a model. Hence, we primarily used the factor loadings (i.e., low loadings) from items loading on factor five as a guide for dropping items to improve the fit. As we dropped items and re-ran analyses, the various fit indexes improved

incrementally. The final model consisted of 35 items in 7 factors with each factor being measured by 4–6 items and all items loading on the a priori factors identified in the exploratory factor analysis.

Final fit statistics were all adequate (with one exception) as follows: χ^2 ($df = 539, N = 297$) = 1,579.57, $p < .001$, $\chi^2/df = 2.93$, CFI = .86, RMSEA = .08, Standardized Root Mean Residual (SRMR) = .06, and 90% confidence interval = .076 – .085. The RMSEA is often considered one of the most valuable fit indices in SEM but should not be used as a universal cutoff rule (Chen, Curran, Bollen, Kirby, & Paxton, 2008). Hu and Bentler (1999) indicated a value of .06 would suggest a good fit, and small RMSEA confidence intervals indicate potential model fit in the population. We also examined the standardized Root Mean Residual (SRMR) where a value below .05 is considered a good fit, and less than .10 is considered favorable (Kline, 2005, p. 141). The CFI (.86) fell short of the typical GOF criteria of .90. However, according to Marsh, Hau, and Wen (2004):

> . . . it is almost impossible to get an acceptable fit (e.g., CFI, RNI, TLI > .9; RMSEA, .05) for even "good" multifactor rating instruments when analyses are done at the item level and there are multiple factors (e.g., 5–10), each measured with a reasonable number of items (e.g., at least 5–10/per scale) so that there are at least 50 items overall. (p. 325)

Given that we started with a 53-item, 7-factor model with 4 to 20 items per factor, our results are consistent with Marsh et al.'s (2004) observations. Clearly, we did not want to engage in the dubious practice of dropping multiple (e.g., 40) items, sacrificing construct validity simply to obtain adequate GOF indices. Like Marsh et al. (2004), we would argue that using GOF indices as rigid decision rules is unwise and not a viable alternative to good judgment based on a sound evaluation of all of the results (i.e., factor loadings, factor correlations, variance accounted for).

Individual factor loadings were all high and ranged from .57 to .90 and clearly met criteria (.40 to .60) established by Hair, Anderson, Tatham, and Black (1998) to designate low to high factor loadings. The associated squared multiple correlations (SMC) range from explaining 33% to 80% of the variance in the factors. Factor correlations ranged from .46 to .86, with most (16/21) ranging from .50 to .75. These moderate correlations and the variance accounted for support the multidimensionality of the scale and indicate that each subscale measures a unique form of self-efficacy. However, based on factor loadings, some items with lower factor loadings (e.g., .57) were not as strong as other items with much stronger factor loadings (e.g., 88). At the same time, the correlations are not so high as to suggest redundant scales or so low as to refute the theoretical tenants of self-efficacy theory. Finally, internal consistency of each scale was also assessed using Cronbach's alpha and ranged from .77 to .91, all meeting the traditional cut off level of .70 (Cronbach, 1951). The above findings are evidence of internal consistency and both convergent and divergent validity.

Factor descriptions

Factor 1: Efficacy about PE content knowledge

The first factor is composed of five cohesive items focusing on confidence in one's level of knowledge to teach certain kinds of PE content effectively. Initially, eight items, to include a wide

selection of activities, were written for the PETES. Factor analysis reduced this to only five, while maintaining variety. Items deleted included team sports, such as basketball, for which participants had consistently high efficacy at all levels of experience, and dance, for which efficacy was generally low. Items include racquet and fitness activities as well as less-traditional content, such as outdoor recreation and aquatics. Higher scores for this factor indicate higher levels of efficacy about one's content knowledge for teaching PE. These items all clearly fit under NASPE element 2.1: "Demonstrate personal competence in motor skill performance for a variety of physical activities and movement patterns" (NASPE, 2009, p. 11), though the teaching aspect of the items is in line with the spirit of "Standard 3: Planning and Implementation: Physical education teacher candidates plan and implement developmentally appropriate learning experiences aligned with local, state, and national standards to address the diverse needs of all students" (NASPE, p. 13).

Factor 2: Efficacy for applying scientific knowledge in teaching PE

The four items loading on Factor 2 pertain to the understanding and application of concepts from the exercise science disciplines (e.g., exercise physiology, motor development), as well as the use of national content standards in planning and instruction. This knowledge and skill appears to be separated from those in the first factor by their abstract nature or because they are acquired in latter stages of the college career through coursework and field experiences. The first item in this factor (Table 2) is like the items in Factor 1, in that it reflects a combination of teaching skill with implied personal competence. The second item aligns with Elements 1.1: "Describe and apply physiological and biomechanical concepts related to skillful movement, physical activity, and fitness;" and 1.2: "Describe and apply motor learning and psychological/behavioral theory related to skillful movement, physical activity, and fitness" (NASPE, 2009, p. 9). The third item is in the spirit of Standard 3, and the fourth aligns with Element 1.3: "Describe and apply motor development theory and principles related to skillful movement, physical activity, and fitness" (NASPE, p. 10). The items in Factor 2 appear to relate not so much to one another as to an underlying construct summarized by the factor's working title of "Things I learned in college." While disparate in content, they are key elements to nearly all programs that candidates would become acquainted with as they progress through their programs. This is in contrast to Factor 1, which is largely what entering physical education teacher education candidates have in mind as they enter their programs.

Factor 3: Efficacy about accommodating skill level differences

Five items relating to skill level differences among students in PE classes loaded onto Factor 3. Skill analysis underlies all items in this factor. These items reflect beliefs about one's abilities to recognize skill level differences via observation, plan a variety of tasks to accommodate for skill level differences in a class, and modify activities to make them developmentally appropriate. These items all fall within Standards 3: "Planning and Implementation: Physical education teacher candidates plan and implement developmentally appropriate learning experiences aligned with local, state, and national standards to address the diverse needs of all students;" and 4: "Instructional Delivery and Management: Physical education teacher candidates use effective communication and pedagogical skills and strategies to enhance student engagement and learning" (NASPE, 2009, pp. 13–17).

TABLE 2
Statistical Characteristics of Manifest Items used in the CFA Analysis for the PETES

Item Characteristic Abbreviations	*M*	*SD*	*SRW*	*SMC*	*CR*
Factor 1: Efficacy about PE content knowledge					
I know a lot about lifetime/recreational games (such as horseshoes, croquet, disc games, cooperative and challenge activities), and can teach them effectively	5.99	2.38	.70	.49	9.4
I know a lot about racquet/net games such as badminton and tennis, and can teach. them effectively	5.72	2.42	.57	0.32	11.0
I know a lot about outdoor recreation activities (such as camping, canoeing, biking, orienteering), and can teach them effectively.	5.78	2.64	.64	0.41	10.2
I know a lot about swimming and water safety, and can teach them effectively.	5.81	2.63	.60	.36	10.6
I know a lot about fitness and can teach it effectively.	7.72	1.92	.72	.51	9.2
Factor 2: Efficacy for applying scientific knowledge in teaching PE					
I have a good grasp of exercise science concepts (from exercise physiology, biomechanics, motor learning and sport psychology), and can apply them to teaching physical education.	6.34	2.26	.69	.74	10.0
I know how first graders are different from fourth graders physically, cognitively, socially, and emotionally.	7.75	1.98	.59	.35	10.7
I know a lot about fundamental motor skills (manipulative and locomotor) and can teach them effectively.	6.72	2.27	.76	.57	8.9
I know what the NASPE standards are, and can plan and teach toward them.	5.71	2.80	.67	.45	10.0
Factor 3: Efficacy about accommodating skill level differences					
When I watch someone perform a skill, I can see if they are doing it right or what they need to correct.	8.28	1.57	.88	.77	8.2
If someone is having trouble performing a skill, I can tell and show them what to do to get better.	8.22	1.60	.90	.81	7.6
I can plan skill sequences so that tasks go from easier to harder in small steps.	7.87	1.76	.74	.55	10.9
If a drill is too easy for a highly skilled student, I can easily change it to make it more challenging.	8.23	1.69	.77	.60	10.2
If one of my students was having trouble with a drill, I know ways to change it to make it easier for them.	7.82	1.66	.78	.61	10.0
Factor 4: Efficacy about teaching students with special needs					
I know how to include a student with cerebral palsy in a regular PE class.	6.04	2.44	.80	.65	9.3
I know what to do with a student with mental retardation in my PE class.	6.13	2.50	.85	.72	8.8
I know how to effectively teach students with emotional or behavioral problems who are in my PE class	7.22	1.22	.87	.76	8.3
If I had a student with vision problems in one of my PE classes, I can find ways for the student to participate with the rest of the class successfully.	7.34	2.14	.73	.53	11.0
I know how to effectively teach a student with ADHD (attention deficit hyperactivity disorder) in my PE class.	7.54	2.05	.86	.73	8.7
Factor 5: Efficacy about instruction					
I can organize and run active classes safely so that students are not likely to get hurt.	8.49	1.43	.76	.58	10.8
I can get my students to respect and cooperate with each other.	8.20	1.47	.74	.55	11.0
I can demonstrate and explain a skill/drill so that the class understands what to do.	8.63	1.34	.80	.65	10.4

(*Continued*)

TABLE 2
(*Continued*)

Item Characteristic Abbreviations	*M*	*SD*	*SRW*	*SMC*	*CR*
I can use clear teaching cues that help students remember and understand how to do a skill correctly.	8.15	1.55	.82	.67	10.3
I can use questions or activities to get kids to think critically or solve problems.	8.04	1.66	.80	.65	10.4
I am able to help children from poverty backgrounds have a successful PE experience.	8.32	1.74	.80	.65	10.4
Factor 6: Efficacy for using assessment					
I understand assessment concepts (such as validity, reliability and authentic assessment) and can apply them in teaching PE.	7.644	1.92	.83	.69	9.2
I can make up rubrics to assess student learning of skills or game play.	7.83	1.81	.84	.71	9.4
My grades reflect how well students have learned what I wanted them to learn.	8.18	1.63	.66	.43	11.1
I can use assessments both for grading my classes and to help me plan.	8.05	1.75	.85	.72	8.8
I can change a lesson as the day goes on based on how the lesson is working.	8.43	1.43	.74	.54	10.5
Factor 7: Efficacy for using technology					
If my principal wants to see me use technology such as computer programs or audiovisual equipment in PE, I can do it.	8.25	1.98	.81	.66	9.1
I can integrate technology if I have access to it (such as video and sound systems) into my teaching.	8.32	1.77	.83	.70	8.4
I am aware of technology-based equipment and computer programs for PE, even if I don't have it.	7.62	2.27	.78	.61	9.4
I often use email and the internet to find or share ideas about PE.	8.08	1.63	.66	.43	11.1
I can use the internet to help plan lessons.	8.32	1.75	.68	.46	10.8

Note: M = Mean. SD = Standard Deviation. SRW = Standardized Regression Weights. SMC = Squared Multiple Correlation. CR = Critical Ratio. SRW, SMC, and CR are from the CFA of the measurement model.

Factor 4: Efficacy for teaching students with special needs

The fourth factor reflects respondents' efficacy beliefs for effectively working with students with special needs in a regular PE class. Several items that related to special needs/differentiating instruction, such as teaching children who were obese, spoke no English, or were poor, did not load in the final PETES. Five items loaded onto this factor. They include a range of cognitive and behavioral diagnoses and included estimations of one's ability to teach these students effectively and modify activities to encourage participation. Higher efficacy scores are also likely to reflect a heightened level of knowledge about particular disorders and grasp of specific strategies for accommodating them. As in Factor 3, NASPE Standards 3 and 4 are the most aligned, with a broad emphasis on effective planning and instructional delivery, including references throughout their elements on meeting the needs of all students.

Factor 5: Efficacy about instruction

Six items represent this factor that reflects efficacy beliefs for what many would consider the act of teaching. Two items reflect each of the three major functions: (1) management—using routines, effective use of space and equipment, managing behavior, (2) motivation—encouraging appropriate levels of class participation and an appreciation of PE, and (3) communication—especially in ways that demonstrate respect and sensitivity to diversity. Higher scores on this factor should be interpreted to reflect greater levels of efficacy to engage in effective in-class teaching behaviors, or what might be considered general components of effective teaching. All items in Factor 5 would align with Standard 4, as they refer to actions in class.

Factor 6: Efficacy for using assessment

Five items, addressing multiple aspects of assessment, loaded onto Factor 6. Items include statements pertaining to understanding and using measurement concepts in teaching, creating rubrics, using assessments for modifying ongoing lessons and planning future ones, and using assessment in grading. Higher scores on this factor reflect stronger teacher beliefs in one's understanding of assessment and effective implementation of it. The five items fall within Standard 5: "Impact on Student Learning:

Physical education teacher candidates utilize assessments and reflection to foster student learning and inform instructional decisions" (NASPE, 2009, p. 18). They reflect an understanding that assessment is more than giving grades, including planning and adapting instruction, and multiple forms of assessment.

Factor 7: Efficacy for using technology

Five items loaded onto the seventh and final factor, all pertaining to the use of technology in planning, teaching, and professional communication. Higher scores on this factor reflect increased levels of efficacy for using technology in planning and teaching and an awareness of available software and hardware that may be using in teaching PE. These align with element 3.7: "Demonstrate knowledge of current technology by planning and implementing learning

experiences that require students to appropriately use technology to meet lesson objectives" (NASPE, 2009, p. 15).

DISCUSSION

Teaching and learning comprise a complex phenomenon, and researchers have found that teacher efficacy plays a significant role. Teachers' sense of efficacy appears to affect basic beliefs about students and instruction and choices of instructional methods and also influences their students' beliefs about their capabilities and learning. As research on teacher efficacy has evolved, scholars have increasingly sought ways to study efficacy in a manner that is more consistent with Bandura's conceptualizations, leaning toward efficacy measures that are specific to subject matter and focusing on specific components of the teaching process (e.g., Martin & Kulinna, 2003; Skaalvik & Skaalvik, 2007; Woolfolk-Hoy & Spero, 2005).

We sought to construct a teacher efficacy instrument specific to teaching PE and used the NASPE Initial Physical Education Teacher Education Standards as a basis for generating ideas/concepts that would be considered important to address. The resulting PETES is a 35-item survey composed of 7 efficacy factors: (a) content knowledge, (b) applying scientific knowledge, (c) accommodating skill level differences, (d) teaching students with special needs, (e) instruction, (f) assessment, and (g) using technology. The resultant factors (areas of efficacy beliefs) align relatively well with the ideas expressed in NASPE Standards and also parallel multifactor efficacy instruments developed for general classroom teaching or other subject areas (e.g., Duncan & Ricketts, 2008; Skaalvik & Skaalvik, 2007; Woolfolk-Hoy & Spero, 2005). Factors are composed of items that logically relate and show high levels of internal consistency. Test–retest reliability indicates stable scores when the survey is repeated over a short time span.

These preliminary results suggest that the PETES is an appropriate instrument for measuring PE teaching efficacy, and it is offered as a tool for studying the development of efficacy and its impact on teacher behavior and student outcomes. However, instrument development and the establishment of validity and reliability is an ongoing process. Hence, researchers are encouraged to continue to evaluate the psychometric properties of the PETES. For instance, the multitrait and multimethod approach is an excellent way to establish validity (Campbell & Fiske, 1959). In keeping with suggestions from other scholars, important areas of study include changes in efficacy beliefs over the teacher preparation program and across the career, factors that affect PE teaching efficacy beliefs, and comparisons of efficacy beliefs to actual teaching performance. Sources of efficacy (Klassen et al., 2011) also remains a significant and incompletely understood issue. Further research might also address differences in efficacy in more and less successful teachers, whether efficacy predicts professional longevity, the relationship between efficacy and student attitudes and learning, and whether graduates of traditional and alternative certification programs differ in their efficacy.

ACKNOWLEDGMENTS

The authors thank Russ Carson of Louisiana State University, Ben Dyson of the University of Memphis, Nell Faucette of the University of South Florida, Cliffa Foster of University of

Mary Hardin-Baylor, Karen Fredenburg of Baylor University, Jon Gray of Sam Houston State University, Jacalyn Lund of Georgia State University, Lynn Owens of Montana State University, Paul Rukavina of Adelphi University, and Hans van der Mars of Arizona State University, for their invaluable assistance in data collection.

REFERENCES

Anderson, R. N., Greene, M. L., & Loewen, P. S. (1988). Relationships among teachers' and students' thinking skills, sense of efficacy, and student achievement. *Alberta Journal of Educational Research*, *34*, 148–165.

Arbuckle, J. L. (2003). *Amos 5.0 users guide*. Chicago, IL: SmallWaters.

Armor, D., Conry-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A. et al. (1976). *Analysis of the school preferred reading program in selected Los Angeles minority schools*, from http://www.rand.org/pubs/reports/2005/R2007.pdf

Ashton, P. T. (1984). Teacher efficacy: A motivational paradigm for effective teacher education. *Journal of Teacher Education*, *35*, 28–32.

Ashton, P. T., Webb, R. B., & Doda, N. (1983). *A study of teachers' sense of efficacy.* Final report. Gainesville, FL: University of Florida, Contract No. 400-79-0075, National Institute of Education.

Baker, P. H. (2005). Managing student behavior: How ready are teachers to meet the challenge? *American Secondary Education*, *33*, 51–64.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Feeeman.

Bandura, A. (2006). Guide to the construction of self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Adolescence and education: Vol. 5: Self-efficacy beliefs of adolescents* (pp. 307–337). Greenwich, CT: Information Age Publishing.

Berman, P., McLaughlin, M., Bass-Golod, G., Pauly, E., & Zellman, G. (1977). *Federal programs supporting educational change: Vol. VII: Factors affecting implementation and continuation*, from http://www.rand.org/pubs/reports/2005/R1589.7.pdf

Bong, M. (2006). Asking the right question: How confident are you that you could successfully perform these tasks? In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 287–305). Greenwich, CT: Information Age Publishing.

Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education*, *16*, 239–253.

Byrne, B. M. (2006). *Structural equation modeling with EQS: Basic concepts, applications and programming*. Mahwah, NJ: Lawrence Erlbaum.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105.

Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, *44*, 473–490.

Centers for Disease Control and Prevention. (2012). *Overweight and obesity*, from http://www.cdc.gov/obesity/data/facts.html

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods & Research*, *36*, 462–494.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 9296–9334.

Dembo, M. H., & Gibson, S. (1985). Teachers' sense of efficacy: An important factor in school improvement. *Elementary School Journal*, *86*, 173–184.

Duncan, D. W., & Ricketts, J. C. (2008). Total program efficacy: A comparison of traditionally and alternatively certified agriculture teachers. *Journal of Agricultural Education*, *49*, 38–46.

Ghaith, G., & Yaghi, H. (1997). Relationships among experience, teacher efficacy, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education*, *13*, 451–458.

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology*, *76*, 569–582.

Glickman, C. D., & Tamashiro, R. T. (1982). A comparison of first-year, fifth-year, and former teachers on efficacy, ego development, and problem solving. *Psychology in the Schools*, *19*, 558–562.

Graham, S., Harris, K. R., Fink, B., & MacArthur, C. A. (2001). Teacher efficacy in writing: A construct validation with primary grade teachers. *Scientific Studies of Reading*, *5*, 177–202.

Guskey, T. R., & Passaro, P. D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal*, *31*, 627–643.

Hair, J. F., Anderson, R. E., Tatham, R, L., & Black, W. C. (1998). *Multivariate data analysis with readings* (5th ed.). Englewood Cliffs, NJ: Prentice Hall.

Henson, R. K. (2002). From adolescent angst to adulthood: Substantive implications and measurement dilemmas in the development of teacher efficacy research. *Educational Psychologist*, *37*, 137–150.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structural analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

Imants, J., & Van Zoelen, A. (1995). Teachers' sickness absence in primary schools, school climate and teachers' sense of efficacy. *School Organisation*, *15*, 77–86.

Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998–2009: Signs of progress or unfulfilled promise? *Educational Psychology Review*, *23*, 21–43.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed). New York, NY: Guilford Press.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers of overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*, 320–341.

Martin, J. J., & Hodges-Kulinna, P. H. (2004). Self-efficacy theory and the theory of planned behavior: Teaching physically active PE classes. *Research Quarterly for Exercise and Sport*, *75*, 288–297.

Martin, J. J., & Kulinna, P. H. (2003). The development of a physical education teachers' physical activity self-efficacy instrument. *Journal of Teaching in Physical Education*, *22*, 219–232.

Martin, J. J., & Kulinna, P. H. (2005). A social cognitive perspective of physical activity related behavior in physical education. *Journal of Teaching in Physical Education*, *24*, 265–281.

Martin, J. J., McCaughtry, N., Hodges-Kulinna., P. H., & Cothran. D. (2008). The influences of professional development on teachers' self-efficacy toward educational change. *Physical Education and Sport Pedagogy*, *13*, 171–190.

Martin, J. J., McCaughtry, N., Kulinna, P. H., Cothran, D., & Faust, R. (2008). The effectiveness of a mentoring-based professional development on physical education teachers' pedometer and computer efficacy and anxiety. *Journal of Teaching in Physical Education*, *27*, 68–82.

National Association for Sport and Physical Education (NASPE). (2001). *Initial physical education teacher education standards* (5th ed.). Reston, VA: Author.

National Association for Sport and Physical Education (NASPE). (2009). *National standards & guidelines for physical education teacher education*. Reston, VA: Author.

Pajares, F., Hartley, J., & Valiante, G. (2001). Response format in writing self-efficacy assessment: Greater discrimination increases prediction. *Measurement and Evaluation in Counseling and Development*, *33*, 214–221.

Riggs, I. M., & Enochs, L. G. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education*, *74*, 625–637.

Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology*, *99*, 611–625.

Soodak, L. C., & Podell, D. M. (1996). Teacher efficacy: Toward the understanding of a multi-faceted construct. *Teaching and Teacher Education*, *12*, 401–411.

Stein, M. K., & Wang, M. C. (1988). Teacher development and school improvement: The process of change. *Teaching and Teacher Education*, *4*, 171–187.

Tournaki, N., & Podell, D. M. (2005). The impact of student characteristics and teacher efficacy on teachers' predictions of student success. *Teaching and Teacher Education*, *21*, 299–314.

Trentham, L., Silvern, S., & Brogdon, R. (1985). Teacher efficacy and teacher competency ratings. *Psychology in the Schools*, *22*, 343–352.

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, *68*, 202–248.

Webb, R. B., & Ashton, P. T. (1987). Teacher motivation and the conditions of teaching: A call for ecological reform. In S. Walker & L. Barton (Eds.), *Changing policies, changing teachers: New directions for schooling?* (pp. 22–40). Philadelphia, PA: Open University Press.

Wheatley, K. F. (2005). The case for reconceptualizing teacher efficacy research. *Teaching and Teacher Education*, *21*, 747–766.

Woolfolk-Hoy, A., & Spero, R. B. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education*, *21*, 343–356.