Original Article

Systematic analysis of physical education standards, benchmarks and related teacher decisions

BUNS MATTHEW T.

Department of Kinesiology and Health Sciences, Concordia University—St. Paul, U.S.A.

Published online: June 26, 2015

(Accepted for publication May 10, 2015)

DOI:10.7752/jpes.2015.02042;

Abstract:

This study examined 16 school district curricula and surveyed teachers in those districts to determine what factors influenced their decisions about what to teach and their professional development. The Physical Education Curriculum Analysis Tool (PECAT) instrument was used to assess each school district's written physical education standards and benchmarks. PECAT content coverage scores ranged from 0-64% ($m=35.2\%$), seven districts had scores of 44% or more. The curriculum map analysis of 1060 benchmarks produced 27% fully aligned, 52% partially aligned and 21% autonomous benchmarks. Five districts had at least one fully aligned benchmark for content associated with their district standards. PECAT and the curriculum maps scores were correlated ($r=.58$, $p=.0001$). Teachers reported little to no professional development related to curriculum. The results of this study partially support Chen's (2006) observation of a "disturbing" misalignment of standards for physical education offered in some schools today. This study goes beyond previous work by providing empirical evidence about curriculum alignment. Better district curriculum maps and PECAT scores were associated with teacher decisions based on student assessment ($e.s.=.86$) and student needs ($e.s.=.81$) by using PECAT and curriculum mapping.

Key words: PECAT, curriculum map, vertical alignment.

Introduction

Quality physical education, as defined in the Physical Education Curriculum Analysis Tool (PECAT), has four components (Centers for Disease Control and Prevention, 2006). One of those is a meaningful curriculum based on standards. Quality physical education has been a key strategy to increase physical activity and reduce health risk. There are two levels of implementation for standards-based curriculum; the standards and benchmarks developed or adopted by the state or district, and the teacher who will use the standards and benchmarks to guide instruction. Standards-based instruction, specifically in physical education, represents a paradigm shift for schools and teachers (National Association for Sport and Physical Education, 2004).

Physical education has a set of competencies that define the skills and knowledge that students are expected to learn through physical education (National Association for Sport and Physical Education, 2004). These competencies have been labeled *content standards*, and specify what a student should know and be able to do as a result of participating in a quality physical education program. While the national content standards describe what students are expected to know and be able to do, they do not define what is considered acceptable performance. That is the role of performance standards, otherwise referred to as *benchmarks*. Benchmarks are specific skills and knowledge that represent progress toward the standards. The revised national content standards (SHAPE America, 2014) are:

- Standard 1 - The physically literate individual demonstrates competency in a variety of motor skills and movement patterns.
- Standard 2 - The physically literate individual applies knowledge of concepts, principles, strategies and tactics related to movement and performance.
- Standard 3 - The physically literate individual demonstrates the knowledge and skills to achieve and maintain a health-enhancing level of physical activity and fitness.
- Standard 4 - The physically literate individual exhibits responsible personal and social behavior that respects self and others.
- Standard 5 - The physically literate individual recognizes the value of physical activity for health, enjoyment, challenge, self-expression and/or social interaction.

Standards-based program design is a process for designing educational programs that begins at the end and works to the beginning of the curriculum. States have control of their own educational system, each state is responsible for developing its own standards. National Association of Sport and Physical Education (NASPE)

has provided national leadership for developing K-12 physical education standards (1995, 2004). The NASPE Standards were based on the document *Outcomes of Quality Physical Education* (National Association of Sport and Physical Education, 1992), which defined a physically educated person. Although each state is responsible for developing its own content standards, many states have adopted or adapted the NASPE National Physical Education Content Standards (2004).

Two paradigms of learning have had an effect on educational program design. The first, *objectives-based program design* (Tyler, 1949) stems from behavioral theory as a foundation of learning, whereas the second, the *standards-based program design*, derived from cognitive theory as a foundation of learning. The objectives-based approach to program design has shaped how teachers contemplate planning and implementing educational programs for decades. Tyler's (1949) process entailed devising a systematic plan for creating content-driven educational experiences through written objectives indicating the behavior that the student would develop. Tyler's work is seminal because it led to the dominant curriculum design process of forward mapping of curriculum, designing educational programs from the bottom up (K-12). Traditionally, programs evolve from bottom up—grade-by-grade, adding forward from the most basic, elemental components at the lower grades to more complex applications at the higher grades. The forward mapping approach created systematic goals and objectives that led to rather narrow and homogeneous learning expectations (Lambert, 2003). Kirk's (1993) analysis of the objectives-based approach to program design identified key limitations of the objectives-based process: (a) objectives lead to "compartmentalization, marginalization, trivialization" of qualitative subjective, and humanistic experiences, and (b) the assumption that motor learning can be easily assessed and measured because of its overt, performance nature and that other forms of learning are not as important.

The standards based approach has fewer components connected across the program, represents the "big ideas," the concepts and principles, not just facts and single elements. This perspective emphasizes what students should know and be able to do when they exit high school. The standards-based program design process is often termed "backward design" or "reverse mapping" because the process leads to programs that are designed from the end back toward the beginning. A primary goal of standards-based program design is to let the standards guide learning (Lambert, 1999).

The process of curriculum mapping (Jacobs, 1997) allows teachers, schools and school districts to examine their physical education curriculum for the content and assessments that they deliver each month over the school year. They align benchmarks to the curriculum map to identify any redundancies across grade levels or any instructional gaps that would reduce students' chances of meeting required benchmarks. The goal has been to address each standard systematically across the curriculum in a sequence that demonstrates a logical and developmentally appropriate progression.

Curricular alignment is expressed in two directions—vertical and horizontal (Thomas, Lee, & Thomas, 2008). Vertical alignment describes the relationship of the benchmarks and content across grades and begins with the mapping of the curriculum. The CDC defines curriculum as a sequential system for delivering meaningful content (Center for Disease Control and Prevention, 2006). The focus on appropriate sequencing includes both developmentally appropriate assessments and ensuring that basic skills lead to more advanced skills. In order to demonstrate sequencing, content must have cohesive threads or skills from grade-to-grade. In most school districts, more than one teacher provides physical education, so implementation of vertical alignment is the result of a plan that is developed and then executed by more than one educator. Vertical alignment of standards and benchmarks may be done at the state, district, or building level, so, in many cases, a physical education teacher does not create this part of the plan. In these cases, it is critical that each physical education teacher accepts responsibility for their portion of the plan.

The Centers for Disease Control and Prevention (CDC) developed the *Physical Education Curriculum Analysis Tool* (PECAT) to assist educators in analyzing the strengths and weaknesses of written physical education curricula in terms of content, student assessment and sequence. PECAT is also used to assess alignment of curricula with national standards, guidelines and best practices for quality physical education programs. Finally, PECAT includes guidance on how to improve curriculum based upon the results.

The benchmarks presented in PECAT describe what children should know and do at the end of grades 2, 5, 8, and 12. PECAT provides 4 to 6 benchmarks for each standard at each of the four levels (grades K-2, 3-5, 5-8, and 9-12). Examination of benchmarks for one standard across the grades should provide evidence of progress toward the standard; however expectations become increasingly more difficult or complex and more similar to the standard as children get older. This concept is consistent with developmentally appropriate physical education. The PECAT score represents the breadth of a curriculum (e.g., number of standards covered) and the depth (e.g., number of benchmarks covered).

Developing a standards-based curriculum begins by looking at the standards, recognizing the skills, knowledge, and dispositions that students should demonstrate to meet these standards, and selecting a curriculum model and/or activities that will allow students to reach the outcomes stated in the standards (Lund, 2005). Since time is limited, teachers must carefully choose content and activities that will allow students to reach the standards. Some activities may be eliminated from a program because of their minimal contribution to meeting standards.

Those developing curricula must decide what they are going to accept as evidence that students have met standards (Lund, 2005; Siedentop, 2005). Additionally, they must decide at what point(s) students are going to demonstrate competence. The process of curriculum planning and assessment should occur simultaneously (Huba & Freed, 1999). Therefore, teachers should have been part of the process from the beginning.

Standards-based curricula represent a paradigm shift for many teachers currently in the field (Doolittle, 2003). It forces teachers to select activities and justify their contribution to meeting the standards rather than selecting activities by teacher preference or tradition. Some students experience a thoughtful variety of activities, with sufficient time and progression in each activity to allow them to achieve the NASPE standards. Other students experience a variety of activities organized with little concern for program goals, and insufficient time in any of the activities to become proficient. The result may be that the students do not meet any NASPE standards.

Little is known about teachers understanding of and attitudes toward the physical education standards. Recent articles (Peterson, Cruzet & Amundson, 2002; Veal, Campbell, Johnson, & McKethan, 2002) have indicated positive results from moving to a standards-based approach in physical education, although these lack empirical evidence. For example, as a result of increased emphasis on standards and accountability measures, authors argue administrators have convinced physical education teachers of the need for, and importance of, standards-based instruction and assessment (Peterson et al., 2002). Administrators have also advocated for resources that will allow teachers to revise curricula and bring programs in line with standards, and for the first time in many districts, teachers were designing specific performance indicators and assessments for how to measure achievement of standards and benchmarks (Veal et al., 2002). States vary widely in how standards and benchmarks are developed; for example Iowa has no state standards and leaves all decisions at the local district level, while Texas has state mandated standards and benchmarks for each grade. Regardless of the origin of the standards and benchmarks, teachers must understand and be able to use the standards and benchmarks for students to meet the standards.

In order for the standards to become guidelines for curriculum, teaching, and assessment, it is important for teachers to gain a keen understanding of the standards. In order to help students achieve desired learning outcomes, it is the teacher's responsibility for embracing and integrating the standards into their daily teaching practices (Fullan, 2001; Glisan, 1996; Leinwand, 1992; Ravitch, 1992). The theoretical framework of teacher change (Fullan, 2001; Glisan, 1996; Leinwand, 1992) suggests teacher's knowledge about and attitudes toward educational standards, their personal commitment to learning about the standards, and their availability and participation in formal professional development activities influence change in teacher's beliefs, knowledge, and behaviors (Fullan, 2001). Thus, standards and the teachers' use of those standards are important to quality physical education.

Chen (2006) investigated the current levels of teacher's knowledge and views of the NASPE standards and factors that influenced the teacher's understandings and interpretations of the standards. Through 25 formal interviews and 78 lesson observations, findings indicated that: a) personal commitment is a key factor contributing to teachers growing knowledge about the standards, b) active participation in professional development activities helps teachers stay current, and c) understanding of the standards is an influential determinant of the teacher's attitude toward the standards.

Articulation of the curriculum across grade levels is a primary concern when implementing standards-based education. There is evidence of a disturbing misalignment between the standards and actual curriculum offered in some schools (Chen, 2006). The main purpose of this project was to assess the alignment of physical education curricula of sixteen school districts where there were no state standards or benchmarks. A secondary purpose was to examine factors underlying curricular decisions by teachers in light of the district curriculum map.

Method

Research Design and Participants

This study was an analysis of physical education standards and benchmarks collected from 16 independent public school districts in one Midwestern state. District administrators were asked to provide copies of their district standards and benchmarks for physical education. All districts were located in one state that does not provide state physical education standards or benchmarks. Initially sixteen districts were invited to participate by letter to the superintendent. The districts represented all geographic regions of the state, and varied on other characteristics for example, larger and smaller enrollment, free and reduced price lunch eligibility, and racial diversity. When a district declined to participate another similar district was invited. The participation rate was 15%, with a total of 110 invited to reach sixteen participants. Standards and benchmarks were required by the state but created at the district level. Some teachers (n=43) from each district volunteered to complete a survey. The study was approved by the Institutional Review Board and all participants completed an informed consent.

Measures

Physical Education Curriculum Analysis Tool (PECAT). The PECAT instrument was used to assess each school district's written physical education standards and benchmarks (Centers for Disease Control and Prevention, 2006). PECAT is a content analysis scoring system rating the written curriculum on each of the six national standards for physical education. The content analyses were divided into four subsections corresponding to the grade-level ranges used in the national standards for physical education: K-2, 3-5, 6-8, and 9-12. Each subsection began with a list of what students were expected to achieve by the end of the identified grade-level range related to each of the national standards. PECAT examines curriculum within grade levels and does not track content across grade levels. A percent of content coverage was calculated using the PECAT protocol. In cases where districts did not have a standard similar to the content of the NASPE standard or did not have benchmarks for any grade within a level the coverage score was zero.

Vertical Alignment. A curriculum map of each district was created tracking cohesiveness of benchmarks across levels (e.g., grades). Reliability for the vertical alignment analysis was previously demonstrated by two trained researchers with agreement over 90% (Thomas, Smith & Buns, 2010). Vertical alignment of district standards and benchmarks was determined by categorizing each series of benchmarks as one of three types based on "tracking" criteria (See Table 1 for examples). Fully tracking sequences refer to each series of benchmarks that tracks across all levels (K-12) for a given standard. Partial tracking sequences were those that tracked across at least two levels (e.g., K-6) but not across all levels (K-12). Autonomous benchmarks were characterized as those that are present once in the entire curriculum.

Table 1. Examples from participating districts to demonstrate definitions of tracking of benchmarks across grades for vertical alignment of NASPE "Standard 1. Students will demonstrate competency in many movement forms and proficiency in a few movement forms".

K-2 Benchmark	3-5 Benchmark	6-8 Benchmark	9-12 Benchmark	Type of Tracking
Use simple combinations of fundamental movement skills	Use mature form and appropriate sequence in combinations of fundamental locomotor object control and rhythmical skills that are components of selected modified games, sports and dances	Perform a variety of simple rhythmic dances	Demonstrate competency with a variety of complex social dance forms	Full*
Use control in weight-bearing and balance activities on a variety of body parts	Use mature form in balance activities on a variety of apparatuses			Part*
	Use basic sport-specific skills for a variety of physical activities	Use beginning strategies for net and invasion	Participate in sporting activities with consistency all of the basic skills, rules and strategies.	Part*
Use a variety of basic object control skills				Autonomous **

Note: * indicates acceptable sequencing; **indicates unacceptable sequencing

Procedures

A trained researcher gathered all district standards and benchmarks and followed established PECAT protocols (CDC, 2006). Curriculum maps were completed and shared with district teachers who volunteered to participate in the project. Teachers reviewed and approved the curriculum map for their district. All data was collected during the spring semester of the academic school year. Teachers completed a survey to identify professional development during the previous year and factors that influenced curricular decisions.

Design and Analysis.

This was a descriptive study. Dependent variables were PECAT percent coverage and rating of benchmarks (full, partial or autonomous). Using the PECAT and vertical alignment scores, districts were placed

in one of two groups representing better and poorer curricula. Teacher survey data was reported based upon their district grouping.

Results

The number of standards for physical education among districts ranged from three-to- seven standards ($M = 5.6$, $SD = 1.3$). Two of sixteen districts used the 2004 NASPE National Content Standards (NNCS) verbatim or with minor modifications, five used a previous version of the NASPE standards (where there were seven standards) and nine created their own standards. All districts included a standard for skill similar to NASPE standard 1, being physically active (NASPE standard 3), and fitness (NASPE standard 4). Valuing physical activity (NASPE standard 6) was the most frequently omitted content at the local level. Districts divided grades into groups in four ways; clusters (e.g., K-2, 3-5, 6-8, 9-12) similar to the PECAT levels, individual grades (e.g., K, 1, 2, 3, 4, 5, 6, 7-8), early end (benchmarks for elementary grades only), and late start (no benchmarks for grades k-3 or k-6). Districts most frequently ($n=8$) divided grades into three-to-five levels. Three districts did not separate benchmarks by grade level and one district provided no benchmarks. The number of benchmarks ranged from zero to 247 ($M=62.5$, $SD=67.6$).

PECAT Content Analysis

PECAT analysis produces a maximum score of 240 “points”; a score of 240 meant that all content was covered at all four PECAT grade groups. In this study, districts addressed 35.2% of PECAT content (Table 2) or 84 of 240 points ($M=84.4$, $SD=58.6$). The benchmarks most thoroughly addressed critical content of NASPE Standard 1 (46.4%) and Standard 4 (43.1%) while Standard 6 was covered at a lower rate (15.0%).

Grade level (elementary, middle school, or high school) PECAT analysis showed that high schools (grades 9-12) most closely aligned their curricula with national standards ($M = 23.9$, $SD = 17.9$, content coverage = 39.8%) when compared with other grade levels (grades 7-8 $M = 22.0$, $SD = 18.1$, content coverage = 36.7%; grades K-6 $M = 42.8$, $SD = 31.3$, content coverage = 35.6%). District enrollment was not related to PECAT total points ($r=-.16$, $p=.56$), or total number of standards ($r=-.33$, $p=.24$).

Table 2. Descriptive statistics for raw scores and PECAT percent coverage by grade level and overall from analysis of 16 school districts

NASPE Standard	PECAT Percent Coverage for Each Grade Level				
	K-2	3-5	6-8	9-12	Overall (K-12)
	Raw Score (<i>SD</i>) (%)	Raw Score (<i>SD</i>) (%)	Raw Score (<i>SD</i>) (%)	Raw Score (<i>SD</i>) (%)	Raw Score (<i>SD</i>) (%)
Standard 1	78.0 ± 22.3 48.8	67.0 ± 19.5 41.9	65.0 ± 21.8 40.6	68.0 ± 21.6 42.5	69.5 ± 21.3 46.4
Standard 4	65.0 ± 31.4 40.6	68.0 ± 34.0 42.5	71.0 ± 33.5 44.4	72.0 ± 33.3 45.0	69.0 ± 33.1 43.1
Standard 2	61.0 ± 29.5 38.1	61.0 ± 32.1 38.1	60.0 ± 31.6 37.5	67.0 ± 31.4 41.9	62.3 ± 31.2 38.9
Standard 3	41.0 ± 28.0 25.6	51.0 ± 30.6 31.9	59.0 ± 30.1 36.9	69.0 ± 29.9 43.1	55.0 ± 29.7 34.4
Standard 5	42.0 ± 29.4 26.3	40.0 ± 31.9 25.0	52.0 ± 31.5 32.5	59.0 ± 31.3 36.9	48.3 ± 31.0 30.2
Standard 6	22.0 ± 25.7 13.8%	21.0 ± 28.3 13.1	24.0 ± 27.8 15.0	29.0 ± 27.6 18.1	24.0 ± 27.4 15.0
TOTAL	51.5 ± 27.7 32.2%	51.3 ± 29.4 32.1%	55.2 ± 29.3 34.2%	60.7 ± 29.2 39.8 %	54.7 ± 28.9 35.2 %

The sixteen districts varied in PECAT coverage scores from 0-64% overall coverage across standards and grade levels (Table 3). Nine districts were at or below 45% coverage and seven districts were above 50% coverage. Of those nine lowest districts, one had no benchmarks and four had one set of benchmarks for all grades k-12. The average number of standards for the districts below 44% and greater than 44% on PECAT coverage was equal at 5.6 standards. Using the same groupings the average numbers of benchmarks were 107 and 18 respectively for the two groups of eight districts. Districts with lower PECAT coverage (below 44%) averaged two grade levels, while the average for higher PECAT districts was five grade levels.

Table 3. Descriptive information by district including source of the standards, number of standards, number of benchmarks, number of grade levels with benchmarks, number of benchmarks labeled as fully aligned, partially aligned and autonomous, PECAT percent coverage score, and number of teachers in the district completing the survey.

ID	Standard source	Number standards	Number benchmarks	Number of grade levels	Fully aligned	Partially aligned	Autonomous	Percent of fully aligned standards	PECAT percent coverage	Teachers completing survey
1	District	4	176	8	4	32	18	75%	64%	3
12	District	5	91	4	8	12	10	100%	64%	2
6	NASPE '95	7	63	4	3	18	3	43%	64%	3
4	NASPE '95	7	247	8	6	41	16	43%	63%	2
2	District	5	82	3	18	8	19	100%	52%	5
3	District	5	80	5	5	16	6	100%	52%	3
5	NASPE '95	7	66	3	11	7	8	100%	52%	5
8	District	5	54	3	4	9	19	100%	45%	3
15	NASPE 2004	6	60	6	6	9	1	67%	32%	1
13	NASPE '95	7	73	4	0	16	6	0%	26	3
16	District	6	26	4	1	3	13	0%	16%	2
10	District	6	16	1	0	0	16	0%	13	1
7	District	3	12	1	0	0	12	0%	10	3
14	District	4	8	1	0	0	8	0%	7	1
9	NASPE '95	7	6	1	0	0	6	0%	3	2
11	NASPE '04	6	0	0	0	0	0	0%	0	3

Vertical Alignment

Curriculum maps were used to assess vertical alignment; this was an analysis of the sequence of benchmarks across grades or levels based on the district standards. Some of the districts used various versions of the NASPE standards, but most used their own standards. Of the 1060 benchmarks in sixteen curriculum maps, 27% were fully aligned, 52% were partially aligned and 21% were autonomous (Table 3). The curriculum maps identified vertical alignment in five districts where at least one benchmark at each grade level focused on related content for each of the district standards. Those districts had five-to-seven standards. Six districts had no benchmarks that were vertically aligned; one had no benchmarks and four others had one grade level.

PECAT does not examine vertical alignment directly, therefore the district benchmarks were matched based on the curriculum maps. Table 4 displays an overall summary of vertical alignment by NASPE standard. *Full* sequencing of benchmarks was the most frequently observed ($M=17.9$, $SD=9.2$) accounting for 27.1% benchmark sequences. *Autonomous benchmarks* were less common ($M=13.8$, $SD=7.8$), accounting for 20.9% of all benchmarks. Larger districts had more benchmarks than smaller districts, however enrollment was not statistically related to vertical alignment ($r=.02$, $p=.94$). More benchmarks were present in the lower levels/grades.

Table 4. Descriptive data from vertical alignment by NASPE standard and overall for 16 school districts including number of full, part, and autonomous benchmarks.

Type of Alignment	Vertical Alignment by NASPE Standard												overall	%
	Standard 1		Standard 2		Standard 3		Standard 4		Standard 5		Standard 6			
	#	%	#	%	#	%	#	%	#	%	#	%		
Full	74	24.0%	39	20.6%	36	28.1%	53	23.9%	57	48.7%	28	29.2%	287	27.1%
Part	159	51.6%	104	55.0%	50	39.1%	121	54.5%	38	32.5%	53	55.2%	552	52.0%
Autonomous	48	15.6%	46	24.3%	42	32.8%	48	21.6%	22	18.8%	15	15.6%	221	20.9%
Total	308	100%	189	100%	128	100%	222	100%	117	100%	96	100%	1060	100%

Note: Full, benchmarks track across all levels (K-12) for a given standard; Part, benchmark sequences track across at least two levels, but not all levels (K-12); Autonomous, a single benchmark is present at one grade level only.

Curriculum Maps and PECAT

PECAT coverage scores and vertical alignment based on the curriculum maps were correlated ($r=.58$, $p=.0001$). Five of the eight highest PECAT coverage scores had fully aligned curriculum maps. Four of the districts deemed fully aligned had PECAT scores above 50%, the other had a PECAT score of 45%. All districts

($n=8$) with PECAT coverage scores above 50% averaged seven fully aligned benchmark sequences (4-18 sequences and an average of 30 benchmarks in the sequences) and averaged 10 autonomous benchmarks. The remaining eight districts with the lowest PECAT coverage scores (0-32%) averaged one fully sequenced benchmark (0-6 sequences and an average of 5 benchmarks in the sequences) and an average of eight autonomous benchmarks.

Teacher Decisions and Professional Development

Forty-three teachers across the 16 districts completed a survey with at least one per district. Twenty three teachers surveyed reported serving on the curriculum development committee in twelve of the districts. Most teachers reported that the district standards (96.2%) the NASPE standards (88.9%) and facilities and equipment (85.2%) had a positive influence on what they taught. All teachers (100%) reported student needs as a positive influence on what they taught. Several variables did not influence what teachers taught, for example pre-service preparation. One variable was a negative influence on what teachers taught; that was parent and community preferences where 22.2% of teachers reported a negative influence (55.6% were neutral).

The PECAT coverage scores were used to place districts into one of two groups. The best ($n=8$) and poorest ($n=8$) standards based on PECAT coverage score. Thus, the answers of the teachers were compared based on the quality of the district standards and benchmarks. There were 14 potential influences in the teacher survey, 7 produce essentially the same response for districts regardless of the quality of the standards and benchmarks (e.g. small effect sizes and overlap of the upper and lower confidence intervals). Those factors included district standards, pre-service preparation, parent and/or community preferences, preparing students for the next grade, training to perform the activity, training to teach the activity, and instructional time. Six produced effect sizes favoring the districts with better standards and benchmarks. Those teachers reported making decisions based on NASPE standards ($e.s.=.30$), a textbook or other instructional materials ($e.s.=.57$), professional development ($e.s.=.58$), the local school wellness policy ($e.s.=.60$), student's needs ($e.s.=.81$) and classroom assessment results ($e.s.=.86$). In addition teachers were asked about professional development activities within the previous year. Six areas of professional development topics (state content standards, national content standards, alignment of instruction to standards, individual differences in student learning, and use of technology to support student learning) did not differ when comparing better to poorer curriculum groups. Professional development focused on assessment was moderately higher in districts with poorer standards and benchmarks ($e.s.= -.50$). There was no clear pattern for teachers serving on the curriculum development committee for better and poorer district curricula whether considering surveyed teachers represented (five of 8 and seven of 8 respectively) or the average number of teachers ($m=2$ and $m=1.9$ respectively) on the committee.

A meaningful standards based curriculum has been included as an indicator of quality physical education (Centers for Disease Control and Prevention, 2006). One goal of quality physical education has been to increase physical activity and improve public health. The quality of the curriculum and the teacher's implementation of the curriculum have been identified as problematic in the literature (Chen, 2006). This study examined 16 school district curricula and surveyed teachers in those districts to determine what factors influenced their decisions about what to teach and their professional development.

NASPE has provided national leadership for developing K-12 physical education standards (1995, 2004). Although each state is responsible for developing its own content standards, many states have adopted the NASPE National Physical Education Content Standards (1995), as did the International Council for Health, Physical Education, Recreation, Sport, and Dance (ICHPERSD), an international physical education organization. The districts in this study did not have a state physical education curriculum with either standards or benchmarks as these decisions are made at the local level. This provided a unique window into curriculum development at the local level where the curriculum was delivered. This system has the potential advantage of greater teacher buy-in because the teachers were likely to participate in the development of the curriculum when compared to situations where the state mandates a curriculum. However, a potential challenge was whether or not local districts have the resources in terms of time and knowledge to develop a well-designed and meaningful curriculum.

PECAT

Curriculum has been characterized by meaningful content that is appropriately sequenced. The content coverage scores are based on the notion that districts will have benchmarks across all grade levels (k-12) with equal weighting at each level. Therefore districts with fewer than four levels would seem to be at a disadvantage. The districts in this study with at least four grade levels averaged 45% coverage (16-64%). Three districts had three grade levels with an average of 50% coverage (45-52%). These three districts demonstrated more depth of coverage than the four districts with four levels. Four of the districts had one grade level, the average coverage in those districts was 8%. The percent coverage for that one level was 33% when not scored against all levels (e.g., against 60 rather than 240 points). To optimize PECAT coverage, having at least three or more levels was helpful. There was no clear relationship between district created standards, the 1995 NASPE or 2004 NASPE standards and PECAT score. This is likely because all districts, regardless of the source of the standards covered skill, physical activity and fitness.

The number of benchmarks was related to higher PECAT coverage scores. The six lowest PECAT coverage scores had 26 or fewer benchmarks. While the PECAT coverage scores above 50% all had over 60 benchmarks. Calculating percent coverage of critical content is one way of clarifying gaps that exist in physical education curriculum. Considering that these districts had no state standards or benchmarks to use as a model for their curricula and many did not use the national standards, PECAT coverage was surprising. Clearly, PECAT was designed in a way that accommodated standards other than the NASPE standards.

The PECAT coverage of standards increased across grade levels for all NASPE standards except standard 1 which focuses on motor skills. Districts may perceive skill as the basis for later activities and therefore focus on skill in the early grades and emphasizing other factors such as tactics in later grades. The coverage for the remaining NASPE standard increased across grade levels. Districts may have had more benchmarks at later grades because of perceptions of student readiness in later grades for the target competencies. Alternatively the fitness, being activity, responsible and so forth may have had a higher priority in the upper grades than in the lower grades. It is also possible that a practical reason such as number of teachers or minutes per week of physical education may have contributed to the increases in coverage. There were not simply more benchmarks at the upper levels. Equal numbers of districts had no benchmarks in elementary grades as those with no benchmarks for high school. It is beyond the scope of this study to know what caused the trend. The increases were small but consistent.

Measuring standards 5 and 6 may have been a barrier in developing benchmarks. Huba and Freed (1999) suggested that learning outcomes, that are the same as standards in this situation, should be developed with both instruction and evaluation (measurement) in mind. Valid and reliable measures of valuing (standard 6) and responsibility (standard 5) may have been a barrier and therefore explain the weakness in district benchmarks and resulting PECAT scores. The mechanisms underlying district selection of standards and benchmarks remain unknown.

These districts would benefit from examining the PECAT coverage scores to increase coverage of critical physical education content. At this time it is not known what content coverage would be appropriate to achieve the standards and to assure a physically active population.

Vertical Alignment

Curriculum develop was a challenge for the districts in this study because it was a local responsibility. Five district physical education curricula had one fully aligned benchmark for each of their standards. Across all districts 27% of the benchmarks will fully aligned. These benchmarks map a path to achieving the standard and define how good is “good enough” (Lund & Tannehill, 2010) for the students. What remains unknown is the definition of “good enough” when we consider articulation of the curriculum vertically. Expecting all benchmarks to track across all levels could narrow the curriculum because some content would not be developmentally appropriate at all grade levels. For example, tactics would be introduced in upper elementary grades but not in the lower elementary grades. Tactics would be appropriate for team and individual sports in middle school and likely seen as part of lifetime sports in high school. Therefore, in this study partial alignment, where a benchmark related to benchmarks at adjacent grade levels but not across all levels was not viewed as poor mapping. Only 21% of all benchmarks were autonomous. These are benchmarks without a clear connection to any previous or later benchmark. Some autonomous benchmarks seem to be related to a “pet project” or specific unit. In light of the near absence of vertical curriculum planning time and professional development about curriculum the number of autonomous benchmarks was relatively low. The overall goal of curriculum mapping was to assure that there are no “gaps” in the curriculum so that students will have a reasonable chance to achieve the standards at the end.

Eight districts had more benchmarks fully aligned than autonomous, with appropriate grade levels for four-to-seven standards. In other words half of these districts had acceptable curriculum maps. Three additional districts had six or seven standards and the appropriate number of grade levels. However, only one had more fully aligned benchmarks than autonomous benchmarks. The number of benchmarks was not a key determinant in these three districts. In one district there were 73 benchmarks, none were fully aligned, and 67 were sequence in partial alignment. In this case it was likely that based on the writing style in the benchmarks that three different teachers or groups of teachers wrote each level of the benchmarks. The third district in this group had half of their 26 benchmarks in fully aligned sequences and the rest were autonomous. These districts had clearly attempted to create effective curriculum maps.

Five of the districts did not have a curriculum map, one because there were no benchmarks and the others because benchmarks were presented for only one grade level. Clearly, these curricula are not comprehensive and developmentally appropriate k-12 programs. In addition, some districts had few standards (*one* with three standards; *two* with four standards). Few standards probably meant a narrow curriculum with less opportunity for all students to become physically educated. Nearly one in three of these districts missed the mark for standards, benchmarks and associated levels. These curriculum maps were consistent with Chen’s (2006) findings of “disappointing”. Further, these district maps were not “good enough”.

Curricular Decisions

Teachers in districts with poorer standards and benchmarks reported a moderate amount of professional development focused on assessment, while teachers in districts with better standards and benchmarks reported minor emphasis on assessment. Better standards were associated with a positive (versus neutral) influence of students' needs and classroom assessment for teacher decisions about what to teach. Conversely for teachers in districts with poorer standards and benchmarks those factors tended to have little or no influence. It is unknown whether teachers actually in the two groups of districts actually had different amounts of professional development on assessment. It could be that districts with poor curriculum maps and therefore poor performance standards did more professional development around assessment with a goal of improving. Another explanation might be that in those districts without a culture of planning and assessment, as indicated by poor curriculum maps, teachers perceived more emphasis on assessment in professional development. In either case, clearly some districts would benefit from professional development about assessment and the relationship of benchmarks to assessment and instruction.

PECAT and Curriculum Maps

While there was overlap between district performance on PECAT and the articulation of their standards across grade levels, each provides slightly different information. PECAT provides a picture of coverage within a grade level and the summed coverage of a broad curriculum based on the content defined in the six NASPE standards. Curriculum mapping focuses on the relationship of one grade level to the next grade level and how cohesive the benchmarks are. Curriculum mapping does not judge breadth while PECAT does. Of course a perfect PECAT score would indicate both cohesive sequences and breadth, assuming the PECAT benchmarks are correct. The goal of both methods was to assist districts to improve their programs so no acceptable score has been established. In this study considering both the curriculum map and PECAT half of the districts were acceptable, three more had attempted to use best practices and remaining five districts had little to no evidence of a developmentally appropriate, meaningful or sequential curriculum. While disappointing, that any district would be unable to produce an acceptable physical education curriculum, it is clear that some districts and their teachers are on the right track.

Summary

The results of this study partially support Chen's (2006) observation of a "disturbing" misalignment of standards for physical education offered in some schools today. This study goes beyond previous work by providing empirical evidence about curriculum alignment by using PECAT and curriculum mapping. Agencies (Centers for Disease Control and Prevention, State Departments of Education) and organizations (NASPE) should support physical education curriculum development in four ways:

1. Provide materials with examples of vertical and horizontal alignment that are readily available to districts at little or no cost.
2. Emphasize developmentally appropriate practice.
3. Train state or district personnel focusing on vertical and horizontal alignment of physical education curricula could be provided.
4. Revise national standards and materials (e.g., PECAT) with particular attention to standards 5 and 6. Particularly how to instruct and measure outcomes (benchmarks) related to these standards.

Professional development and vertical curriculum meetings for physical educators in the district with a focus on vertical alignment may be one avenue to improve physical educator knowledge of the standards and benchmarks (Chen, 2006) and the articulation of district standards and benchmarks.

References

- Centers for Disease Control and Prevention. (2006). *Physical Education Curriculum Analysis Tool*. Atlanta, Georgia.
- Chen, W. (2006). Teachers' knowledge about and vies of the national standards for physical education. *Journal of Teaching in Physical Education*, 25, 120-142.
- Doolittle, S. (2003). *Assessment programs in New York State: A whole village's effort*. Paper presented at the National Conference of the American Alliance for Health, Physical Education, Recreation, and Dance, Philadelphia, PA.
- Fullan, M.G. (2001). *The new meaning of educational change* (3rd Ed.). New York: Teachers College Press.
- Glisan, E. W. (1996). A collaborative approach to professional development. In R.C. Lafayette (Ed.), *National standards—A catalyst for reform* (pp. 57-95). Lincolnwood, IL: National Textbook.
- Huba, M., & Freed, J.E. (1999). *Learner-centered assessment on college campuses: shifting the focus from teaching to learning* (1st Ed.). Columbus, OH: Allyn and Bacon.
- Jacobs, H.H. (1997). *Mapping the big picture*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Lambert, L.T. (1999). *Standards-based assessment of student learning: A comprehensive approach*. Reston, VA: NASPE.

-
- Lambert, L.T. (2003). Standards-based program design: Creating a congruent guide for student learning. In S. J. Silverman & C. D. Ennis (2nd Ed.), *Student learning in physical education* (pp 129-144). Champaign, IL: Human Kinetics.
- Leinwand, S.J. (1992). Sharing, supporting, risk taking: First steps to instructional reform. *The Mathematics Teacher*, 85, 466-470.
- Lund, J. & Tannehill, D. (2005). Introduction to standards-based curriculum development. *Standards-based physical education curriculum development* (pp. 3-15). Sudbury, MA: Jones and Bartlett Publishers.
- Lund, J. & Tannehill, D. (2010). Introduction to standards-based curriculum development. In J. Lund and D. Tannehill *Standards-based Physical Education Curriculum Development*, second edition. Sudbury, MA: Jones and Bartlett Publishers.
- Kirk, D. (1993). Curriculum work in physical education: Beyond the objectives approach? *Journal of Teaching in Physical Education*, 12, 244-265.
- National Association for Sport and Physical Education. (2004). *Moving into the future: National physical education standards*. Reston, VA: American Alliance for Health, Physical Education, Recreation and Dance.
- National Association for Sport and Physical Education. (1995). *Moving into the future: National physical education standards*. Reston, VA: American Alliance for Health, Physical Education, Recreation and Dance.
- National Association for Sport and Physical Education. (1992). *Outcomes of quality physical education programs*. Reston, VA: Author.
- Peterson, S., Cruz, L., & Amundson, R. (2002). The standards impact on physical education in New York State. *Journal Physical Education, Recreation, and Dance*, 73 (4), 15-23.
- Ravitch, D. (1992). National standards and curriculum reform: A view from the Department of Education. *NASSP Bulletin*, 76(548), 24-29.
- SHAPE America. (2014). *National standards & grade-level outcomes for K-12 physical education*. Human Kinetics.
- Siedentop, D. (2005). Foreword. In J. Lund & D. Tannehill (Eds.), *Standards-based physical education curriculum development* (pp. xiii-xv). Sudbury, MA: Jones and Bartlett Publishers.
- Thomas, K.T., Smith J.D., & Buns, M.(2010). Analysis of standards and benchmarks from 14 school districts. *Research Quarterly for Exercise and Sport*, 81 supplement, A-74.
- Thomas, K.T., Lee, A.M., & Thomas, J.R. (2008). *Physical Education Methods for Elementary Teachers* (3rd Ed.). Champaign, IL: Human Kinetics.
- Tyler, R.W. (1949). *Basic principles of curriculum and instruction*. Chicago: University of Chicago Press.
- Veal, M.L., Campbell, M., Johnson, D., & McKethan, R. (2002). The North Carolina PEPSE Project. *Journal of Physical Education, Recreation and Dance*, 73 (4), 19-23.