The Physical Educator Vol. 75 • pp. 341–373 • 2018
https://doi.org/10.18666/TPE-2018-V75-I3-8113

ASSESSMENT

Systematic Observation of Formal Assessment of Students by Teachers (SOFAST)

Hans van der Mars, Gay Timken, Jeff McNamee

Abstract

Assessment is a central function of the teaching–learning cycle and a key performance standard for beginning and experienced teachers. Especially in today's school climate of high-stakes teacher evaluation, physical education will continue to be at risk unless programs can begin to provide evidence of student learning. The Systematic Observation of Formal Assessment of Students by Teachers (SOFAST) instrument was developed for the purpose of collecting process data on various dimensions of teachers' formal and informal assessment practices. SOFAST is a three-level observation system that allows users to collect data on teachers' primary teaching functions (including formal and informal assessment), the focus of their assessment, and common contextual dimensions of lessons. Seven secondary-level licensed physical education teachers were observed over three regular lessons. Percentages were calculated for each SOFAST category. Individual and group means and standard deviations were calculated. Secondary physical education teachers employ mostly informal assessment strategies and limit their formal assessment to student efforts on managerial aspects of performance.

Hans van der Mars is a professor and the Physical Education Program director, Mary Lou Fulton Teachers College, Arizona State University. Gay Timken is a professor, Division Health and Exercise Science, Western Oregon University. Jeff McNamee is chair and professor of Human Performance, Department of Health, Human Performance, and Athletics, Linfield College. Please send author correspondence to hans.vandermars@asu.edu

Acknowledgment: This research was funded in part by the Research Consortium of the American Alliance for Health, Physical Education, Recreation, and Dance (AAHPERD).

Assessment for learning (also referred to as formative assessment and learning-oriented assessment; López-Pastor et al., 2013) in physical education is part of a larger educational and assessment movement toward promoting the use of standards-based assessment in the context of standards-based instruction (Lund & Veal, 2013). Increasingly, schools are required to demonstrate that their students are meeting state and national content standards in academic subjects such as science and math. Along with national physical education content standards, the Society of Health and Physical Educators (SHAPE America) has invested heavily in designing assessment resources and tools that evaluate student attainment of these standards (e.g., Lambert, 2007; Lund, 2000; Mitchell & Oslin, 1999; SHAPE America, 2015a, 2015b).

According to SHAPE America (2016), only about 76% of U.S. states require high school physical education to be provided for graduation, and since 2012 the percentage of U.S. junior high and middle schools requiring physical education has decreased (from 84% to 76%). Moreover, less than a third of states require student assessment that is related directly to state physical education standards (SHAPE America, 2016). Coupled with the recognition that their school's administration, school district, and state government have no explicit expectations for student learning in physical education, physical educators are less likely to assess student learning and performance formally compared to their classroom colleagues in other subjects (e.g., Tousignant & Siedentop, 1983).

However, the emergence of high-stakes teacher evaluation schemes in the form of value-added models makes assessment of student outcomes in physical education an even more critical teaching function. School district teacher evaluation models generally include the components of (a) teaching process evaluation, (b) student achievement data, and (c) years of experience, of which the first two make up the largest percentage of the teacher evaluation. Typically, school administrators in districts using value-added models evaluate the teaching process using generic teacher evaluation tools (e.g., Danielson, 2007; Marzano Research Laboratory, 2013). Despite documented inherent problems associated with high-stakes teacher evaluation (e.g., American Educational Research Association, 2015; Berliner, 2014; Lavigne, 2014), as of 2015, 43 states now

require objective measures of student achievement to be included in teacher evaluations, and student growth is the preponderant criterion in teacher evaluations in 16 states (National Council on Teacher Quality, 2015).

Within this culture of high-stakes teacher evaluation, Rink (2007) argued that more than ever it is imperative that physical education programs can (a) clearly articulate the intended program outcomes and (b) provide evidence to parents, school administrators, and policy makers that students are learning something substantive. Logically, this places ongoing formative and formal assessment of student performance and learning squarely as a central teaching function for physical educators. Throughout this paper, the terms *learning*, *outcomes*, and *performance* are used interchangeably, as learning is an ongoing process that occurs throughout units of instruction, as opposed to a single performance event as demonstrated during traditional end-of-unit skill or knowledge tests.

A significant body of professional literature has been dedicated to promoting formal-formative and informal assessment and serves as evidence of the paradigm shift advocated by physical education teacher educators (e.g., Brands, 1996; Cleland & Stevenson, 1997; Lund & Veal, 2013; Melograno, 2000; Siedentop, 1994; Siedentop, Hastie, & van der Mars, 2011; Smith & Cestaro, 1998; van der Mars & Harvey, 2010; Weinberg, 1996; Wood, 2003). This shift points to the need for aligning or integrating assessment with curriculum and instruction (Guskey, 2003; Lund & Tannehill, 2014; Lund & Veal, 2008; Veal, 1992, 1995). Guskey (2003) noted the importance of teachers coming to view ". . . their assessments as an integral part of the instruction process and as crucial for helping students learn" (p. 6). Formative assessment occurs throughout the course of instruction and can offer a more immediate indication of student progress.

The importance of formal-formative assessment to document learning outcomes and its potential to support and motivate students has been widely accepted in the assessment, pedagogy, and teacher education literature (e.g., Black & Wiliam, 1998a, 1998b; Greenwood & Maheady, 1997; Pryor & Akwesi, 1998; SHAPE America, 2014; Wood, 2003). Reviews of research on how classroom teachers' use of formative assessment affects student learning outcomes produced effect sizes ranging from .20 to .70 standard deviation units

(depending on the meta-analysis protocols used; Black & Wiliam, 1998a; Briggs, Ruiz-Primo, Furtak, Shepard, & Yin, 2012; Dunn & Mulvenon, 2009). Within the limitations of the research, classroom teachers' use of formative assessment positively affects students' use of (meta-)cognitive strategies, and student outcomes to varying degrees (e.g., Baas, Castelijns, Vermeulen, Martens, & Segers, 2015; Furtak et al., 2016; Veldhuis & van den Heuvel-Panhuizen, 2014).

According to Pryor and Akwesi (1998), teachers tend to dislike summative (i.e., end-of-unit) assessments because of the lack of timely information on student learning. If one agrees that the quality and quantity of student engagement in learning tasks is a reflection of their learning, it stands to reason that developing records of student progress in this process is essential. Thus, formally assessing student performance should not be limited to end-of-unit (i.e., summative) assessments, such as a skill or written test. As well, assessment of student learning is a key performance standard for beginning and experienced physical educators (National Association for Sport and Physical Education, 2010; National Board for Professional Teaching Standards, 2014).

Physical Educators' Assessment Practices: The Evidence

Most evidence on physical educators' assessment practices dates from the 1980s–1990s. Data on such practices were collected mostly via surveys, interviews, or document analysis (e.g., Hensley, Lambert, Baumgartner, & Stillwell, 1987; Imwold, Rider, & Johnson, 1982; Kneer, 1986; Lund & Veal, 2008). Physical educators frequently employ informal assessment strategies in the form of (non-)verbal feedback, praise, and/or encouragement (e.g., Graber, 2004; Rink & Mitchell, 2003; Tannehill, van der Mars, & MacPhail, 2015). However, they have also shown reticence toward using formal assessment of any kind, whether for supporting student learning or determining students' grades (e.g., Tannehill et al., 2015; Tousignant & Siedentop, 1983). More recently, Leirhaug and MacPhail (2015) found that Norwegian physical educators' " . . . understanding and enactment of AfL [Assessment for Learning] key strategies was somewhat constricted" (p. 637).

Although physical educators generally perceive assessment as important, they also see it as too time consuming, find little value in it, do not think it necessary, and/or lack the requisite knowledge (e.g.,

Kneer, 1986; Matanin & Tannehill, 1994; Veal, 1988). Historically, physical educators have typically employed subjective judgment of student attitude, participation, sportsmanship, and effort as primary performance indicators (e.g., Hensley et al., 1987; Imwold et al., 1982; López-Pastor et al., 2013; Matanin & Tannehill, 1994). Especially in secondary school physical education programs, students' course grades are based mostly on features of management including attendance, on-time behavior, general conduct, and dress (Tousignant & Siedentop, 1983; Veal, 1988). Such expectations for students are generally the most explicit and most closely monitored compared to student learning outcomes on subject-matter-related tasks. Furthermore, although (formal) assessment of learning outcomes is widely accepted as a central teaching function, evidence suggests that few teachers integrate it into their day-to-day teaching (Shepard, 2001). This supports Veal's (1988) claim of a gap between the theory and practice of assessment in physical education programs. Finally, Lund and Veal (2008) found that preservice physical education teachers have great difficulty planning for assessment.

An ample amount of literature reports on why and how physical educators should employ formal and formative assessment on learning outcomes. However, less is known about their assessment practices. With few exceptions, much of the research on physical educators' assessment practices is dated and focuses primarily on their perceptions and opinions. Today's changing context for physical education programs warrants new research on the assessment practices of physical educators. The Systematic Observation of Formal Assessment of Students by Teachers (SOFAST) instrument was developed as a tool for collecting process data on various dimensions of teachers' formal and informal assessment efforts specific to student learning outcomes. This paper (a) provides an overview of the key features of SOFAST including the various coding levels, category definitions, and observation tactics and procedures and (b) presents an initial descriptive data set based on observations of secondary school physical educators through the use of SOFAST.

SOFAST System Description

SOFAST is a three-level observation system that includes the use of partial interval recording and momentary time sampling (van

der Mars, 1989). In sequence, observers code teachers' teaching functions, the focus of teachers' assessment activities, and the lesson context. The Appendix includes a copy of the coding form.

Level 1: Teaching Functions

At the Teaching Functions coding level, the observer measures teachers' distribution of time across Assessment, Instruction, and Management functions, using partial interval recording with alternating 10-s "observe" and 10-s "record" intervals. The assessment can be either informal or formal (relative to whether a permanent record is created). Formal assessment (coded F) occurs when teachers record information on either paper or an electronic device about their students' performance of physical education content learning tasks, managerial tasks, and general social behavior. Informal assessment (coded I) occurs when teachers provide students with (non-) verbal feedback and encouragement on learning tasks, managerial tasks, and general social behavior.

The Instruction function is encompassed in two codes: the teacher's time spent Participating (or demonstrating; P) with students and his/her time spent sharing Knowledge (K) about physical education content (e.g., how to serve in volleyball). It reflects attention to how and when teachers communicate information about learning tasks. Code K oftentimes occurs when students are inactive (e.g., listening to instruction), but it may also occur when students are engaged in an activity (e.g., 3-v-3 modified volleyball games on four courts) and the teacher is monitoring the activity. If or when the teacher stops one of the games to discuss technical or tactical aspects of game play, the observer would code this as K.

The Managerial function (M) reflects time allocated to organizational activities, such as how and where equipment is to be placed, taking attendance, forming groups, and discussions regarding students' overall class conduct (i.e., personal and social behavior). If the teacher is silently observing the class for the entire interval, the observer would mark this as O.

Level 2: Teacher Assessment Focus

The second level of the decision sequence involves coding for the focus of teachers' assessment efforts, using partial interval recording with alternating 10-s observe and 10-s record intervals. The choice

is fourfold. If, at the first coding level, the teacher was coded as not engaged in either formal or informal assessment, the observer would enter a 0 (i.e., zero). If the teacher was coded as engaging in either formal and/or informal assessment, the observer would code the focus of the assessment.

If the teacher assesses student actions within a physical education motor- or knowledge-related learning task, the observer would code this as C (for Content). If the teacher assesses student performance on managerial tasks (e.g., getting grouped quickly, freezing quickly, or starting a task quickly), the observer would code this as M (for Management). The S (Social Behavior) code would be marked if the teacher focuses the assessment on students' social or personal conduct.

Level 3: Lesson Context

The third and final decision level targets the context of the lesson, with a focus on the activity of the whole class. The observer uses momentary time sampling, taking the sample every 20 s upon the record cue. Observers need to decide between whether lesson time is allocated for management activities (M) or for lesson or unit content. The coding of lesson content might include a decision to determine whether the focus is on knowledge (K) or motor (physical activity) content. If the activity reflects the latter, the observer must decide if the context reflects a focus on fitness (F), skill or technique practice (S), game play (G), or other (O).

SOFAST Data Collection Protocol

SOFAST data are collected using a combination of partial interval recording and momentary time sampling at 20-s intervals using a 10-s observe/10-s record format, for three observation samples per minute. For the Teaching Functions level and Teacher Assessment Focus level, the observer has 10 s to observe and determine the behavior pattern and then 10 s to record that decision. Upon the record cue, the observer also samples the lesson context at that time (i.e., momentary time sampling). Observers use observe and record prompts via prerecorded digital MP3 files.

Level 1: Teaching Functions Category Definitions

The observer uses a coding hierarchy to code the teaching functions. For example, during observe intervals in which multiple behaviors occur (e.g., Formal assessment and Knowledge), the observer codes the one that is listed further left in the Level 1 category order (i.e., F, I, P, K, M, and O). Furthermore, given the variability in the natural duration of the teaching functions, the observer uses partial interval recording. That is, a behavior need not occur for the entire duration of the interval (McKenzie & van der Mars, 2015; van der Mars, 1989).

Formal assessment (F). Using either paper and pen or an electronic data collection device (e.g., tablet), the teacher records information about student performance on tasks that are related to physical education content, management or organization, or personal or social conduct.

Informal assessment (I). The teacher provides (non-)verbal feedback to students about performance on tasks related to physical education content, management or organization, or personal or social conduct, but does not collect data. This would include teachers' responses (either positive or corrective) to students' answers to questions (e.g., checking for understanding).

If formal and informal assessment occur during one observe interval (regardless of the order), the observer would mark both coding symbols in the column for Teaching Functions behaviors. Such instances would be direct examples of teachers' effort at employing formal assessment to communicate to students the progress they are making while learning.

Participates or demonstrates (P). The teacher participates with students in physical education content tasks during activity episodes (e.g., fitness routines, 5-v-5 Ultimate Frisbee) and/or demonstrates a skill or technique at times students are to observe the demonstration. This may also include live demonstrations performed by students and/or the showing of a short video clip.

Knowledge (K). The teacher instructs or communicates verbally, including verbal prompts, about physical education content (e.g., technique execution, tactical features of game play, rules, history, officiating, coaching, scorekeeping, judging a dance routine,

heart rate checks). This might include the use of a whiteboard, video, and/or printed graphics.

Management (M). Time spent by the teacher on organizational tasks (that are or can be "routines") unrelated to lesson content in the form of general directions, including prompts. This refers to the explanation and practice of tasks such as moving equipment on or off the field or floor, forming small student groups, dispersing students to activity stations, gathering students together, and moving students from one area to another, such as from the gym to an outdoor field. It also includes time spent on recognizing positive or inappropriate negative student conduct (e.g., treatment of equipment, facilities, talking during class-wide instruction, safety-related behavior).

Silent observation (O). The teacher monitors student performance without interacting with them during periods of subject-matter activity. Silent observation may also include talking to oneself, unintelligible statements (e.g., incomplete sentences), and engaging in conversation with a person who is not part of the class activities (e.g., student intern, principal, other teachers).

Level 2: Teacher Assessment Focus Category Definitions

As was the case with the Teaching Functions level, the observer uses partial interval recording to code the assessment focus. Teachers may assess multiple dimensions of student performance (e.g., Content and Management) during one observe interval. In this case, the observer would record the focus that lasted the longest during that interval.

0 (zero). No informal or formal assessment of student learning outcomes occurred. This would be coded if the Teaching Functions behavior code for that interval was Participation, Knowledge, Management, or Silent Observation.

Content—Motor and knowledge focus (C). Formal or informal assessment by teachers that targets learners' performance on physical education knowledge and/or motor content tasks. Examples include performance in fundamental motor skills, health or fitness activities, the technical execution of exercises in strength conditioning, sport-related techniques and tactics, choreography of dance steps, quality of team play, and/or knowledge of an activity (e.g., a correct student response to a teacher's question about a rule, tactic, strategy, historic event, famous athletes). Assessment pertaining to

non-playing student roles in Sport Education (e.g., Game Official, Coach, Judge, Scorekeeper, Fitness Trainer, Sport Board Member; Siedentop et al., 2011) would also be coded as being focused on content.

Management focus (M). Formal or informal assessment by teachers that targets learners' managerial and organizational task (i.e., routines) performance. Examples include assessment of students' grouping, gathering and dispersing of students and equipment, students' quick or slow start to a learning task, students' responses to freeze signals, hustles, compliance with attendance procedures, being on time, and being dressed correctly for class. Teachers' assessment comments about students' general class safety behaviors would also be coded as having a management focus.

Social and personal behavior focus (S). Formal or informal assessment by teachers of students' social and/or personal conduct in class. Examples include assessment of how students cooperate with each other or with the teacher, how they interact and treat each other, how they include peers in an activity, and how they assist peers who experience difficulty with a task or activity.

The social and personal behavior focus would also include assessment of students being attentive during class-wide instruction, making supportive comments to others, and demonstrating a liking for an activity or showing that they value an activity.

Recognizing aspects such as hard work, perceived effort, and/or persistence (e.g., during fitness activities) would be coded with S. The observer would mark this code when teachers give feedback to students about taunting, swearing, bullying, (mis-)treating equipment and/or facility, and so forth. Assessment during lessons that have a direct and clear focus on individual or social skill development in students (e.g., team-building content) would also be coded as S.

Level 3: Lesson Context (adapted from McKenzie, 2015)

For the Lesson Context level, observers use momentary time sampling. That is, the observer records the predominant activity that the class as a whole is engaged in upon the record cue. Thus, given the 20-s interval length, this produces three observation samples per minute.

Management (M). Refers to lesson time when students are not intended to be involved in physical education content, including transition, management, and break times. This includes time allocated to managerial and organizational activities such as taking attendance, discussing students' conduct, team or group formation, changing or moving equipment, moving from one space to another (e.g., from gym to weight room or from gym to outdoor facility), changing stations (e.g., going from a basketball activity to a pickleball area), and changing activities within a lesson.

This includes teachers' explanations or demonstrations of organizational or managerial arrangements, as well as students' execution of such managerial tasks. If such episodes are followed by physical-education-content-related activities, the management episode ends when the first student starts engaging in the content-related task.

Transitions that occur naturally as part of an activity would be coded as part of that activity rather than as Management (M). For example, time spent moving from one fitness circuit station to the next would be coded as Fitness (F), and changing sides of the court during a volleyball game would be coded as Game (G). However, if or when an activity is halted for more than 10 s, the observer would enter a new code (most likely M or Kn).

Knowledge (Kn). Lesson time in which the primary focus is on student knowledge acquisition related to physical education content and/or cognitive activities such as writing tasks (e.g., journaling); math-related tasks (e.g., entering pedometer step counts); or student groups discussing how to develop a dance, jump rope, or gymnastics routine. Knowledge is related to conveying information about (a) health-related physical activity and fitness content (e.g., physical activity or physical fitness concepts, including endurance, critical elements for executing a bench press, differences between static and ballistic stretching, taking a heart rate, and discussions about making consumer choices regarding fitness products) and (b) skill-related content, such as dance, sport, and outdoor pursuits (e.g., skill cues and common errors for the track and field baton exchange, tactical aspects of game play, how to set a screen in basketball, how to use a compass in orienteering, historic events in sport, using game rules

effectively, etiquette). This knowledge may be conveyed by teachers, other students, guest speakers, video, and computers, among others.

Fitness (F). Class time allocated to activities with a major purpose of altering the physical state of individuals in terms of cardiovascular endurance, strength, and/or flexibility. Examples include aerobic dance, calisthenics, distance running, strength conditioning, agility training, fitness testing, and warm-up and cooldown activities. Relays conducted with three or more students per team would be coded as Game (G).

Skill or technique (S). Class time devoted to students practicing the execution of techniques with the primary goal of technique development (e.g., basketball layup, dance steps, or balance beam skills).

Game (G). Class time devoted to applying techniques, tactics, and strategy in a game form or competitive setting. Game contexts could include either parent versions of sport games (e.g., 5-v-5 basketball on full court with standard rules) or games that include modifications to either primary or secondary rules (e.g., 2-v-2 volleyball; 3-v-3 soccer; altered rules, court size, net height, etc.). Combative challenges (e.g., in wrestling) would be coded as Game (G). Team-building and icebreaker activities, tag games, and judged performances such as gymnastics and dance routines that are part of a class competition would also be coded as G.

Other (O). Refers to free playtime during which physical education instruction is not intended (e.g., time when students choose their own activities and make their own decisions about rules). This time resembles recess, during which students may select to participate or not.

As noted, much of the previous research on teachers' assessment practices was based on either surveys or interviews. The following section presents initial data from observations of secondary school physical educators that were made using SOFAST.

Method

Participants and Settings

Four female and three male certified physical education teachers served as participants. Their teaching experience ranged from 4 to 15

Table 1
Demographics of Participating Schools

Participating school	MS 1	MS 2	MS 3	MS 4	HS
Grades	6–8	6–8	6–8	6–8	9–12
Enrollment (*n*)	1,042	570	1,389	574	1,327
Economic Disadvantage (%)[a]	32	28	12	43	31
Students With Disabilities (%)	14	8	10	12	10
English Language Learners (%)	21	8	20	12	< 5
Ethnicity (%)					
American Indian/Alaska Native	0	1	0	0	1
Asian	9	9	40	1	1
Black/African American	3	1	2	1	1
Hispanic/Latino	20	8	7	19	12
Multi-Racial	6	8	6	6	4
Native Hawaiian/Pacific Islander	1	0	0	0	0
White	60	73	46	74	81
Curricular Focus	Multiactivity	Multiactivity	Multiactivity	Multiactivity	Fitness for Life/Sport Education

Note. Data from State's Department of Education Annual Report Cards. MS = middle school; HS = high school.
[a]Students eligible for free or reduced meals.

years. Teachers were employed at middle schools (n = 6) and a high school (n = 1). Teaching contexts were varied. One middle school was located in an affluent neighborhood of a suburban community that serves as a bedroom community for a large metropolitan city. One middle school was located in a suburban middle-class neighborhood. The remaining two middle schools and one high school were located in more rural communities. Table 1 presents additional school demographic information. Several teachers had a sustained record of professional development, presenting at state and regional conferences and professional organizations. Two participating teachers had earned state-level teacher of the year awards.

The programs in which teachers taught were varied in curricular orientation. In three of the middle school programs, teachers employed the Dynamic physical education framework (Darst & Pangrazi, 2006). Because of space limitations, teachers in the fourth middle school grouped their three classes together (class sizes range: 33 to 45) in one gym for a common fitness activity. After that, students were free to select between multiple activities (typically at least three options) in various parts of campus and switch between them on a daily basis. The high school teacher's curricular focus represented a blend of Fitness for Life (Corbin & Lindsey, 2005) and Sport Education (Siedentop et al., 2011).

Lesson length for all middle schools was 50 min (including dress time), and classes met daily, except for one school, which employed an A-B schedule (i.e., physical education met every other school day). The high school employed an A-B block schedule (i.e., five lessons per 10 school days) with lessons scheduled for 90 min.

Data Collection and Analysis

Each participant was observed teaching three regular physical education lessons. Trained observers used SOFAST. Well-defined behavior categories of systematic observation instruments provide low-inference descriptions of observable behaviors and events. As such, they have strong internal (or face) validity (McKenzie, 2010; McKenzie & van der Mars, 2015). The use of six external reviewers asssessing face validity ensured that the SOFAST category definitions in the first two decision levels (Teaching Functions and Assessment Focus) were defined clearly and allowed for clear discriminations between categories. Reviewers were selected based on their expertise

in the use of systematic observation tactics. Based on reviewer suggestions, minor adjustments were made in three category definitions. Lesson Context definition categories were not checked for validity as they were drawn from well known and extensively used systematic observation systems (e.g., System for Observing Fitness Instruction Time [SOFIT], McKenzie, Sallis, & Nader, 1991; Academic Learning Time-Physical Education [ALT-PE], Parker, 1989).

Percentages were calculated for each SOFAST category. Means and standard deviations were calculated for each teacher across the three observed lessons. These means were then averaged, resulting in group means and standard deviations.

Observer Reliability

Observer reliability was established through interobserver agreement (IOA) checks for at least one of the three observed lessons with each participating teacher, by a second trained and independent observer. IOA percentages were calculated with the total interval method (van der Mars, 1989). As shown in Table 2, IOA percentages for Teaching Functions (including formal and informal assessment), Assessment Focus, and Lesson Context data were at acceptable levels across teachers. Of the 105 individual category IOA percentages, 13% were in the 80–90% range, which were all the result of being low-occurrence behaviors or episodes. All others were at or above 90%. Based on the IOA results, the observers were deemed reliable.

Results

Class Context Data

Figure 1 includes means and standard deviations for the various class activities. Teachers allocated 57.4% (SD = 13.52) of the class time in some form of physical activity. Fitness-related content was the most prominent type of physical activity (M = 25.7, SD = 19.85). This was followed by Game (M = 17.1, SD = 9.31) and Skill Practice (M = 15.5, SD = 11.85), respectively. Time spent in class-wide instruction (coded as Knowledge) averaged at 15.9% (SD = 6.58) of the available class time, whereas managerial activities took up an average of 26.7% (SD = 9.38) of the class time.

Table 2
Total Interval (T-I) Inter-observer Agreement (IOA) Percentages Across Participants.

	Teacher						
Level and category	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Teaching Functions							
Formal Assessment	97.50	100.00	93.75	91.60	98.00	100.00	97.40
Informal Assessment	92.30	88.00	90.15	94.50	90.30	94.70	92.90
Participation/Demonstration	100.00	91.00	95.50	100.00	88.90	90.00	96.10
Knowledge	80.00	68.00	82.00	100.00	96.80	91.60	99.00
Management	90.30	94.00	92.15	98.70	99.00	100.00	100.00
Silent Observation	94.40	98.00	63.70	90.30	93.70	100.00	92.30
Teaching Functions Average	92.42	89.83	86.21	95.85	94.45	96.05	96.28
Assessment Focus							
Content	100.00	86.00	93.00	96.30	90.20	98.70	98.40
Management	100.00	100.00	100.00	95.00	97.90	81.80	90.50
Social Behavior	96.00	100.00	89.00	100.00	93.40	100.00	100.00
Assessment Focus Average	98.67	95.33	94.00	97.10	93.83	93.83	96.30
Class Context							
Management	90.20	94.00	92.10	97.60	94.00	86.30	98.10
Knowledge	100.00	94.00	97.00	94.90	90.60	90.50	93.90
Fitness	100.00	100.00	100.00	100.00	100.00	100.00	95.50
Skill Practice	97.80	96.00	96.90	98.90	98.90	100.00	100.00
Game	96.60	100.00	98.30	92.90	92.90	100.00	94.50
Other	100.00	100.00	100.00	100.00	98.00	100.00	93.90
Class Context Average	97.43	97.33	97.38	97.38	95.73	96.13	95.98

Teaching Functions Data

Figure 2 shows group means and standard deviations on the SOFAST Teaching Functions categories, with the most prominent being Management (M = 30.5%, SD = 6.37) and Instruction (M = 29.1%, SD = 8.34). Instruction was broken down by the percentage of class time spent in Knowledge and Participation/Demonstration, which averaged 22.9% (SD = 5.74) and 6.2% (SD = 5.68), respectively. Teachers engaged in various types of assessment an average of 36.0% (SD = 13.86) of the class time. Teachers allocated an average of 3.9% (SD = 3.64), 20.7% (SD = 15.80), and 0.2% (SD = . 44) of the class time to Formal, Informal, and simultaneous Formal and Informal assessment, respectively. Finally, teachers spent an average of 14.5% (SD = 9.79) of the class time in Silent Observation of students.

Teacher Assessment Focus

Figure 3 includes data on the focus of teachers' assessment. Two thirds of the teachers' assessment effort targeted physical-education-content-related motor and knowledge performance (M = 67.5%, SD = 15.86). The next most common focus of teachers' assessment was student performance on Management-related tasks (M = 25.8%, SD = 15.43). Finally, teachers' assessment of students' Social Behavior made up 6.7% (SD = 9.28) of the assessment.

Teachers' Assessment Patterns

Figure 4 includes a breakdown of teachers' assessment relative to type and focus. When formally assessing students, teachers focused mostly on their managerial task performance (M = 57.4%, SD = 35.79). The remaining formal assessment was split between students' Content-related task performance (M = 8.6%, SD = 19.22) and social or personal behavior (M = 0.7%, SD = 1.80). Most informal assessment was directed to students' Content task performance (M = 79.5%, SD = 15.37). Student performance on Management tasks was assessed informally an average of 13.5% (SD = 9.06) of the informal assessment, whereas the remaining 7.0% (SD = 9.79) targeted students' social or personal conduct.

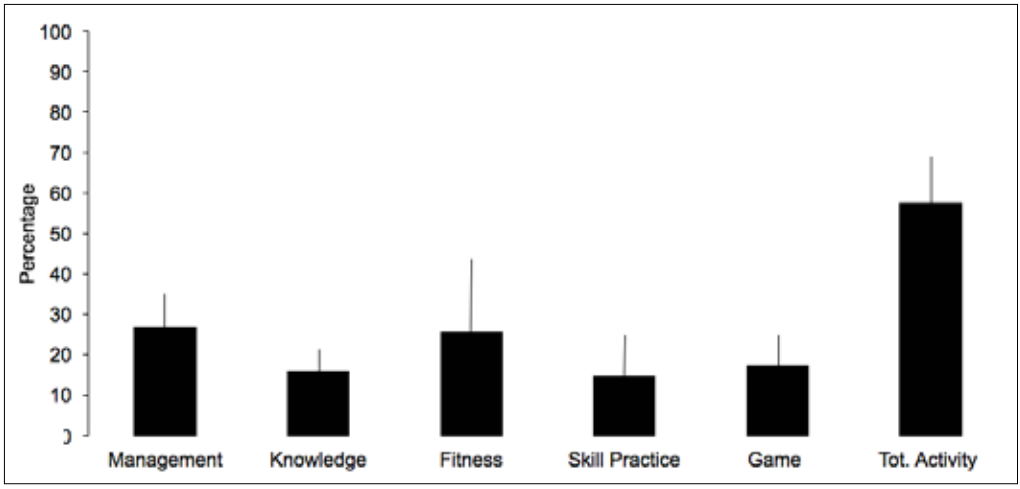

Figure 1. Mean percent of class time for lesson context categories

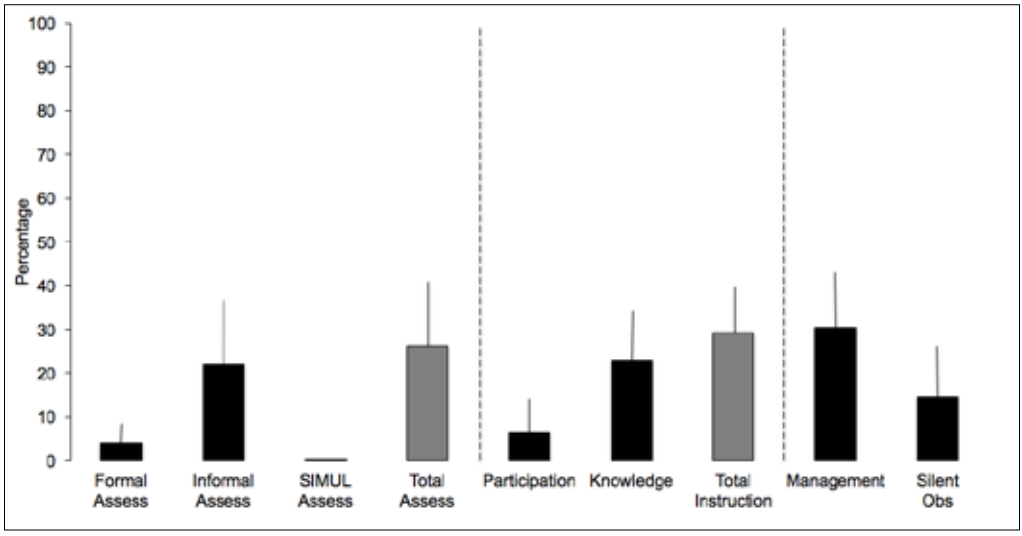

Figure 2. Mean percent of class time across teaching functions.

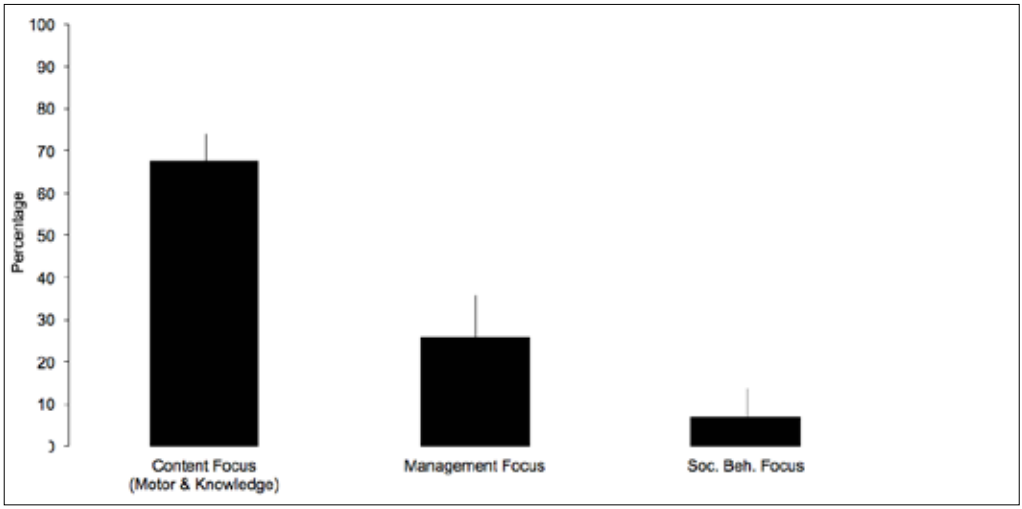

Figure 3. Percent of total assessment by focus.

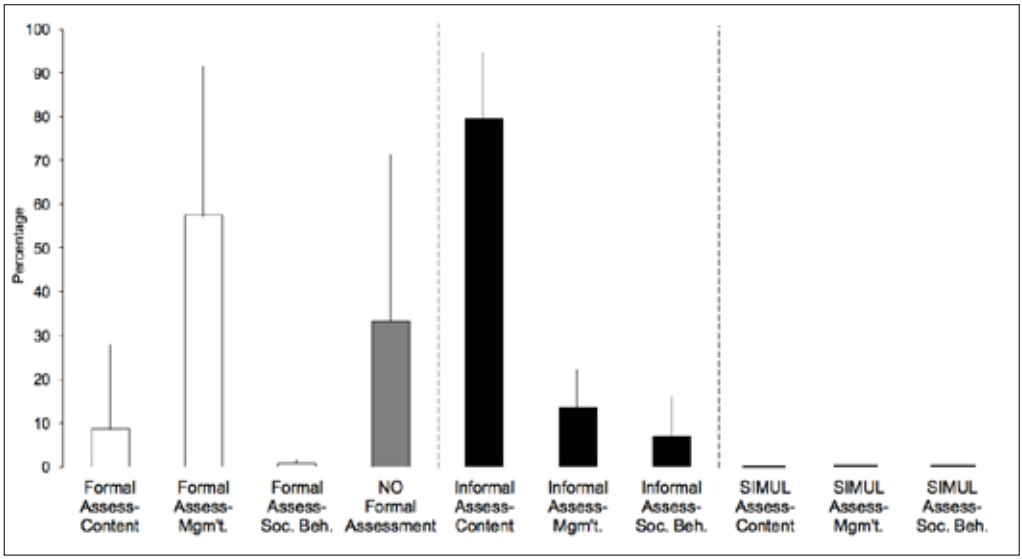

Figure 4. Mean percent of assessment by type.

Discussion

The process data on participating teachers' formal and informal assessment practices confirm findings from most previous studies that physical educators do assess their students (e.g., Graber, 2004; Rink & Mitchell, 2003; Siedentop & Tannehill, 2000). However, their assessments consisted mainly of informal assessment of student performance on content-related tasks, and formal assessment was limited to their students' performance on attendance, dress, and timely arrival to class. This reflects a lack of change in assessment practices reported previously (e.g., Graber, 2004; Rink & Mitchell, 2003; Tannehill et al., 2015; Tousignant & Siedentop, 1983).

The almost 8 min of class time that teachers spent silently observing students is somewhat below the levels reported in previous research (Siedentop, 1991; Siedentop & Tannehill, 2000). The professional physical education teacher education (PETE) literature describes this monitoring task as essential to ensuring that (a) safety is maintained in the learning environment, (b) students make progress on learning tasks, and (c) they conduct themselves appropriately. What remains largely unknown is what teachers think about, observe, and make decisions about during those bouts of silent observation. Are the observations deliberate and focused, or are the teachers using that time to "take a breather"?

Physical educators have voiced that formal assessment is too time consuming and lacks relevance (Goc-Karp & Woods, 2008; Kneer, 1986). This view may be, in part, a consequence of how they conceptualize assessment in general and formal assessment (i.e., summative: skill testing at the end of a unit) in particular. Rink, French, Lee, Solmon, and Lynn (1994) found that student teachers did not include assessment (labeled as *evaluation* in their study) in their descriptions of effective teaching. If teachers were to view formal assessment more as a day-to-day process that occurs throughout units of instruction than as something that is not important or valued, the time spent in silent observation potentially lends itself to not only focused and deliberate observations, but also the recording of data on student performance and progress.

Balancing and integrating the central teaching functions of instruction, management, and monitoring of students for the purpose of creating appropriate student learning experiences in itself

requires a strong command of teaching skills. As noted, current trends in school districts' teacher evaluation practices reflect an increased emphasis on tying teacher evaluations directly to student achievement. The recent passing of the Every Student Succeeds Act (ESSA, 2015) stipulates that states and school districts must develop and implement teacher (and principal!) evaluation systems that are based in part on evidence of student achievement. Thus, formal-formative assessment of student outcomes in physical education should figure more prominently and likely requires adjustments in how teachers attend to other core teaching functions. For example, teachers who are effective in employing managerial structures and routines and who can get students to be more self-managed (e.g., Sport Education) can carve out precious class time that can then be used for formal assessment. Likely, programs in which students largely depend on frequent teacher directions or in which the teacher's classroom management skills are less well developed will leave little if any opportunity for (formal) assessment. In addition, there have been calls in (physical) education for increasing the emphasis on assessment for learning (i.e., formative), wherein data are collected throughout the learning process (e.g., Baker & Gordon, 2014; Hay, 2006; van der Mars & Harvey, 2010), as well as for a de-emphasis of end-of-unit or end-of-year testing.

Physical educators have reported feeling ill-equipped to conduct assessment of their students (Kneer, 1986). Regardless of whether one agrees with school districts' current teacher evaluation practices, physical educators who want to strengthen their case for receiving positive evaluations must be equipped with the needed assessment skills that will help them demonstrate what their students are achieving in physical education. The historic (and continuing!) lack of focus by physical educators on authentic, formal-formative assessment of learning outcomes points to the need for in-depth and sustained professional development support (Leirhaug & MacPhail, 2015).

SHAPE America (and many individual teachers) regard development of appropriate social and personal conduct skills as central program outcomes (e.g., Goc-Karp & Woods, 2008; SHAPE America, 2014), and assessment tools specific to these outcomes are available (e.g., SHAPE America, 2015a, 2015b). In this study, teachers focused surprisingly little of their assessment on these very

dimensions of espoused student learning (i.e., less than 10% of all assessment). Formal assessment of social and personal conduct was virtually absent.

When, during a unit of instruction, teachers assess students formally (i.e., develop a permanent record) and then simultaneously communicate this to the students, this constitutes an explicit example of aligning assessment with instruction so it can support student learning. In none of the observed lessons did any of the teachers use what little formal assessment information they did collect as a springboard for providing their students with feedback or prompts to aid their students' progress, regardless of which aspect of student performance (i.e., content, management, or social or personal).

A strength of this study is that despite the variance in teaching contexts, experience, and curricular orientation among participating teachers, their assessment patterns were largely similar, as evidenced by the relatively small standard deviations in the various Teacher Assessment Focus categories. When assessing formally, teachers focused almost entirely on students' managerial performance, a pattern of assessment which has been seen in previous research (e.g., Hay, 2006; Tousignant & Siedentop, 1983).

This study was not without limitations. First, the sample of participants prevents generalization to all secondary physical educators. However, the findings from this study largely align with previously reported data. Another limitation was that because participating teachers did not focus any of their formal assessments on substantive student learning outcomes, we could not show how the use of the SOFAST observation system can discriminate between teachers who make and do not make formal assessment of student outcomes a primary focus.

The SOFAST observation instrument presented in this paper allows researchers to gain insight into physical educators' use of formal and informal assessment, its focus (i.e., student learning indicators and outcomes, and indicators of student managerial and social behavior performance), and several contextual dimensions of classes. SOFAST category definitions of behavior are low-inference and include numerous examples of behaviors and events, which thus eases the process of memorizing them. As with all systematic observation instruments, reliable and accurate use of the system depends

on the quality of observer training that precedes its use (McKenzie & van der Mars, 2015).

Implications

The presented data must be placed in the larger context of school physical education programs, which continue their struggle to maintain a viable presence in public schools and to maintain relevance and credibility in the eyes of the public. On the one hand, significant losses have occurred in the allocation of minutes per week and in graduation requirements for physical education on school campuses (e.g., Center on Education Policy, 2007; SHAPE America, 2016; U.S. Department of Health and Human Services, 2015). Yet, at the same time, federal agencies, national research bodies, and professional organizations strongly support school physical education programs, addressing the role of schools in promoting students' physical activity. Schools are widely regarded as a critical point of impact in promoting and increasing physical activity opportunities for children and adolescents (e.g., Centers for Disease Control and Prevention, 1997, 2001; Institute of Medicine, 2013; National Physical Activity Plan Alliance, 2016; Pate et al., 2006; Payne & Morrow, 2009). For example, Pate et al. (2006) recommended,

> States should hold schools accountable for delivering PE programs that meet national standards for quality and quantity (i.e., 150 minutes per week for grades K to 8 and 225 minutes per week for grades 9 to 12). Specifically, each state should include PE in its core educational accountability system and should incorporate PE into its system national standards for curriculum and instructional quality. (p. 7)

At the federal level, there have been multiple unsuccessful policy efforts to elevate physical education as a school subject to the same level as classroom "core" subjects (e.g., FIT Kids Act, 2009). In 2015, the U.S. Congress passed the Every Student Succeeds Act (ESSA) that replaced No Child Left Behind. As noted, teacher evaluation that is based in part on student achievement remains in place. However, school health and physical education are school subjects identified as part of students' "well-rounded" education. Moreover, Title IV funding associated with ESSA will flow to states, of which 20% must

fund "safe and healthy schools," and must go toward "well-rounded subject" areas.

Although the ultimate outcomes of such policy and legislative successes remain unknown as yet, this creates state-level opportunities for physical education to move forward. The key will be how K–12 professionals, PETE faculty, and sport pedagogy scholars respond. Both the previous research and the absence of any appreciable amount of formal assessment of student performance on subject-matter-related learning tasks in this study suggest we are not prepared. Theoretically, this lack of preparedness is at least in part a consequence of (a) the absence of state-level accountability mechanisms that require physical education programs to provide evidence that their students are meeting content standards (a notable exception being South Carolina; Rink & Mitchell, 2003); (b) the absence of not only reliable and valid, but also usable and practical assessment tools; (c) how (formal) assessment is conceptualized (Rink et al., 1994); (d) the lack of good training in using assessment of student outcomes within PETE programs (e.g., Lund & Veal, 2008) and in-service professional development opportunities for teachers; and (e) a lack of necessary skills and knowledge about effective assessment among physical educators (e.g., Kneer, 1986; Veal, 1988).

Lund and Veal (2008) noted that the inability of student teachers to (a) develop appropriate learning objectives, (b) select appropriate assessments, and (c) implement assessment of student learning is likely the result of the absence of a clear assessment culture in PETE programs and the profession at large. If school physical education is to lay legitimate claim to part of the school day, it can no longer consider itself immune from having to demonstrate that investments in their programs are warranted. From a perspective of being proactive, PETE programs and any professional development programs for already certified teachers must increase efforts in equipping current and future teachers with the skills, knowledge, and (perhaps most important!) dispositions necessary to make assessment of student learning a part of their daily work. Failure to develop a culture of assessment will likely reinforce students, parents, school administrators, and policy makers to view physical education as an easily disposable school subject and will likely perpetuate the perceived marginalization of the subject.

We have argued that there is an urgent need for new and experienced physical educators to make formal-formative assessment a more integral teaching function in their day-to-day work with students. As noted, much of the previous research on physical educators' use of assessment is dated and lacks a focus on assessment practices. Today's policy context for physical education offers potential, but may leave physical education more vulnerable because of the lack of a strong evidence base. Additional research on assessment practices (and their effect on student outcomes) is warranted. To that end, researchers can use SOFAST to objectively track physical educators' in-class assessment practices. In addition, it can be used to determine the effect of interventions aimed specifically at the development of formal-formative assessment skills. Moreover, supervisors in PETE programs can use SOFAST to track progress in developing such assessment skills among preservice PETE students. A complete description of the complete SOFAST coding manual can be obtained from the lead author.

Conclusion

The following conclusions can be drawn from the current study: First, SOFAST can provide important process data regarding physical education teachers' engagement in formal and informal assessment of student learning and performance. Second, secondary physical education teachers continue to limit their formal assessment to student efforts on managerial aspects of performance.

References

American Educational Research Association. (2015). AERA statement on use of value-added models (VAM) for the evaluation of educators and educator preparation programs. *Educational Researcher, 44,* 448–452. https://doi.org/10.3102/0013189X15618385

Baas, D., Castelijns, J., Vermeulen, M., Martens, R., & Segers, M. (2015). The relation between Assessment for Learning and elementary students' cognitive and metacognitive strategy use. *British Journal of Educational Psychology, 85,* 33–46. https://doi.org/10.1111/bjep.12058

Baker, E. L., & Gordon, E. W. (2014). From the assessment of education to the assessment for education: Policy and futures. *Teachers College Record, 116,* 1–24.

Berliner, D. C. (2014). Exogenous variables and value-added assessments: A fatal flaw. *Teachers College Record, 116,* 1–31.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy, and Practice, 5,* 7–74. https://doi.org/10.1080/0969595980050102

Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80,* 139–144.

Brands, P. (1996). The pros outweigh the cons: An authentic assessment approach to teaching physical education. *Teaching Secondary Physical Education, 2*(4), 18–21.

Briggs, D. C., Ruiz-Primo, M. A., Furtak, E., Shepard, L., & Yin, Y. (2012). Meta-analytic methodology and inferences about the efficacy of formative assessment. *Educational Measurement: Issues and Practice, 31*(4), 13–17. https://doi.org/10.1111/j.1745-3992.2012.00251.x

Center on Education Policy. (2007). *Choices, changes, and challenges: Curriculum and instruction in the NCLB era*. Washington, DC: Author.

Centers for Disease Control and Prevention. (1997). Guidelines for school and community programs to promote lifelong physical activity among young people. *MMWR Recommendations and Reports, 46*(RR-6), 1–37.

Centers for Disease Control and Prevention. (2001). Increasing physical activity: A report on recommendations of the Task Force on Community Preventive Services. *MMWR Recommendations and Reports, 50*(RR18), 1–16.

Cleland, F. E., & Stevenson, R. (1997). Authentic assessment in action. *Teaching Secondary Physical Education, 3*(3), 4–7.

Corbin, C. B., & Lindsey, R. (2005). *Fitness for life* (5th ed.). Champaign, IL: Human Kinetics.

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching.* Alexandria. VA: Association for Supervision and

Curriculum Development. https://doi.org/10.1037/0735-7028.38.6.652

Darst, P. W., & Pangrazi, R. P. (2006). *Dynamic physical education for secondary school students* (5th. ed). San Francisco, CA: Pearson.

Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessments: The limited scientific evidence of the impact of formative assessments in education. *Practical Assessment, Research, and Evaluation, 14*(7), 1–11.

Every Student Succeeds Act, S.1177, 114 Cong. (2015).

FIT Kids Act, S. 634, 111 Cong. (2009).

Furtak, E. M., Kiemer, K., Circi, R. K., Swanson, R., de León, V., Morrison, D., & Heredia, S. C. (2016). Teachers' formative assessment abilities and their relationship to student learning: Findings from a four-year intervention study. *Instructional Science, 44,* 267–291. https://doi.org/10.1007/s11251-016-9371-3

Goc-Karp, G., & Woods, M.L. (2008). Preservice teacher's perceptions about assessment and its implementation. *Journal of Teaching in Physical Education, 27,* 327–346. https://doi.org/10.1123/jtpe.27.3.327

Graber, K. (2004). Research on teaching in physical education. In V. Richardson (Ed.), *Handbook of research on teaching* (pp. 491–519). Washington, DC: American Educational Research Association.

Greenwood, C. R., & Maheady, L. (1997). Measurable change in student performance: Forgotten standard in teacher preparation? *Teacher Education and Special Education, 20,* 265–275. https://doi.org/10.1177/088840649702000307

Guskey, T. R. (2003). How classroom assessments improve learning. *Educational Leadership, 60*(5), 6–11.

Hay, P. J. (2006). Assessment for learning in physical education. In D. Kirk, D. Macdonald, & M. M. O'Sullivan (Eds.), *Handbook of physical education* (pp. 312–325). London, England: Sage. https://doi.org/10.4135/9781848608009.n18

Hensley, L., Lambert, L., Baumgartner, T., & Stillwell, J. (1987). Is evaluation worth the effort? *Journal of Physical Education, Recreation, and Dance, 58*(6), 59–62. https://doi.org/10.1080/07303084.1987.10609577

Imwold, C. H., Rider, R. A., & Johnson, D. J. (1982). The use of evaluations in public schools physical education programs. *Journal of Teaching in Physical Education, 2,* 13–18. https://doi.org/10.1123/jtpe.2.1.13

Institute of Medicine. (2013). *Educating the student body: Taking physical activity and physical education to school*. Washington, DC: National Academies Press.

Kneer, M. (1986). A description of physical education instructional theory/practice gap in selected secondary schools. *Journal of Teaching in Physical Education, 5,* 91–106. https://doi.org/10.1123/jtpe.5.2.91

Lambert, L. (2007). *Assessment Series K–12 Physical Education. Standards-based assessment of student learning: A comprehensive approach*. Reston, VA: National Association for Sport and Physical Education.

Lavigne, A. L. (2014). Exploring the intended and unintended consequences of high-stakes teacher evaluation on schools, teachers, and students. *Teachers College Record, 116*. Retrieved from http://www.tcrecord.org/content.asp?contentid=17294

Leirhaug, P. E., & MacPhail, A. (2015). 'It's the other assessment that is the key': Three Norwegian physical education teachers' engagement (or not) with assessment for learning. *Sport, Education, and Society, 20,* 624–640. https://doi.org/10.1080/13573322.2014.975113

López-Pastor, V. M., Kirk, D., Lorente-Catalán, E., MacPhail, A., Macdonald, D., & Lorente, E. (2013). Alternative assessment in physical education: A review of international literature. *Sport, Education, and Society, 18,* 57–76. https://doi.org/10.1080/13573322.2012.713860

Lund, J. (2000). *Assessment Series K–12 Physical Education. Creating rubrics for physical education*. Reston, VA: National Association for Sport and Physical Education.

Lund, J., & Tannehill, D. (Eds.). (2014). *Standards-based curriculum development* (3rd ed.). Sudbury, MA: Jones and Bartlett.

Lund, J., & Veal, M. L. (2008). Chapter 4: Measuring pupil learning—How do student teachers assess within instructional models? *Journal of Teaching in Physical Education, 27,* 487–511. https://doi.org/10.1123/jtpe.27.4.487

Lund, J., & Veal, M. L. (2013). *Assessment-driven instruction in physical education: A standards-based approach to promoting and documenting learning*. Champaign, IL: Human Kinetics.

Marzano Research Laboratory. (2013). Marzano teacher evaluation model. Retrieved from http://www.marzanoevaluation.com

Matanin, W. C., & Tannehill, D. (1994). Assessment and grading in physical education. *Journal of Teaching in Physical Education, 13,* 395–405. https://doi.org/10.1123/jtpe.13.4.395

McKenzie, T. L. (2010). 2009 C. H. McCloy Lecture Seeing Is Believing: Observing physical activity and its contexts. *Research Quarterly for Exercise and Sport, 81,* 113–122. https://doi.org/10.1080/02701367.2010.10599656

McKenzie, T. L. (2015). *System for Observing Fitness Instruction Time: Description and procedures manual*. Retrieved from http://activelivingresearch.org/sofit-system-observing-fitness-instruction-time

McKenzie, T. L., Sallis, J. F., & Nader, P. R. (1991). SOFIT: System for Observing Fitness Instruction Time. *Journal of Teaching in Physical Education, 11,* 195–205. https://doi.org/10.1123/jtpe.11.2.195

McKenzie, T. L., & van der Mars, H. (2015). Top 10 research questions related to assessing physical activity and its contexts using systematic observation. *Research Quarterly for Exercise and Sport, 86,* 13–29. https://doi.org/10.1080/02701367.2015.991264

Melograno, V. J. (2000). Designing a portfolio system for K–12 physical education: A step-by-step process. *Measurement in Physical Education and Exercise Science, 4,* 97–115. https://doi.org/10.1207/S15327841Mpee0402_5

Mitchell, S., & Oslin, J. (1999). *Assessment Series K–12 Physical Education. Assessment in games teaching.* Reston, VA: National Association for Sport and Physical Education.

National Association for Sport and Physical Education. (2010). *National standards and guidelines for physical education teacher education* (3rd ed.). Reston, VA: Author.

National Board for Professional Teaching Standards. (2014). *Physical education standards for teachers of students ages 3–18+* (2nd ed.). Retrieved from http://boardcertifiedteachers.org/sites/default/files/ECYA-PE.pdf

National Council on Teacher Quality. (2015). *2015 state teacher policy yearbook: National summary*. Retrieved from http://www.nctq.org/dmsView/2015_State_Teacher_Policy_Yearbook_National_Summary_NCTQ_Report

National Physical Activity Plan Alliance. (2016). *National physical activity plan*. Retrieved from http://physicalactivityplan.org/docs/2016NPAP_Finalforwebsite.pdf

Parker, M. (1989). Academic Learning Time-Physical Education (ALT-PE), 1982 revision. In P. W. Darst, D. Zakrajsek, & V. H. Mancini (Eds.), *Analyzing physical education and sport instruction* (pp. 195–206). Champaign, IL: Human Kinetics.

Pate, R. R., Davis, M. G., Robinson, T. N., Stone, E. J., McKenzie, T. L., & Young, J. C. (2006). Promoting physical activity in children and youth: A leadership role for schools—A scientific statement from the American Heart Association Council on Nutrition, Physical Activity, and Metabolism (Physical Activity Committee) in collaboration with the councils on cardiovascular disease in the young and cardiovascular nursing. *Circulation, 114,* 1214–1224. https://doi.org/10.1161/CIRCULATIONAHA.106.177052

Payne, V. G., & Morrow, J. R. (2009). School physical education as a viable change agent to increase youth physical activity. *President's Council on Physical Fitness and Sports Research Digest, 10*(2), 1–8.

Pryor, J., & Akwesi, C. (1998). Assessment in Ghana and England: Putting reform to the test of practice. *Compare, 28,* 263–275. https://doi.org/10.1080/0305792980280304

Rink, J. (2007). PE teaching: It's ALL about outcomes [Editorial]. *pelinks4u, 9*(7). Retrieved from http://www.pelinks4u.org/archives/070107.htm

Rink, J., French, K., Lee, A., Solmon, M., & Lynn, S. (1994). A comparison of pedagogical knowledge structures of preservice students and teacher educators in two institutions. *Journal of Teaching in Physical Education, 13,* 140–162. https://doi.org/10.1123/jtpe.13.2.140

Rink, J., & Mitchell, M. (Eds.). (2003). State level assessment in physical education: The South Carolina experience. *Journal of Teaching in Physical Education, 22*(5), 471–608. https://doi.org/10.1123/jtpe.22.5.471

Shepard, L. A. (2001). The role of classroom assessment in teaching and learning. In V. Richardson (Ed.), *Handbook of research on*

teaching (4th ed., pp. 1066–1101). Washington, DC: American Educational Research Association.

Siedentop, D. (1991). *Developing teaching skills in physical education* (3rd ed.). Mountain View, CA: Mayfield.

Siedentop, D. (1994). Authentic assessment through sport education. In D. Siedentop (Ed.), *Sport education: Quality PE through positive sport experiences* (pp. 115–118). Champaign, IL: Human Kinetics.

Siedentop, D., & Tannehill, D. (2000). *Developing teaching skills in physical education* (4th ed.). Mountain View, CA: Mayfield.

Siedentop, D., Hastie, P., & van der Mars, H. (2011). *Complete guide to Sport Education* (2nd ed.). Champaign, IL: Human Kinetics.

Smith, T. K., & Cestaro, N. G. (1998). *Student-centered physical education: Strategies for developing middle school fitness and skills*. Champaign, IL: Human Kinetics.

Society of Health and Physical Educators. (2014). *National standards and grade-level outcomes for K–12 physical education*. Champaign, IL: Human Kinetics.

Society of Health and Physical Educators. (2015a). *Updated standard 1 assessments for PE metrics ebook: Assessing standards 1–6 in elementary school*. Champaign, IL: Human Kinetics.

Society of Health and Physical Educators. (2015b). *Updated standard 1 assessments for PE metrics ebook: Assessing standards 1–6 in secondary school*. Champaign, IL: Human Kinetics.

Society of Health and Physical Educators. (2016). *Shape of the nation: Status of physical education in the USA*. Retrieved from http://www.shapeamerica.org/advocacy/son/2016/upload/Shape-of-the-Nation-2016_web.pdf

Tannehill, D., van der Mars, H., & MacPhail, A. (2015). *Building effective physical education programs*. Sudbury, MA: Jones and Bartlett.

Tousignant, M., & Siedentop, D. (1983). A qualitative analysis of task structures in required secondary physical education classes. *Journal of Teaching in Physical Education, 3,* 47–57. https://doi.org/10.1123/jtpe.3.1.47

U.S. Department of Health and Human Services. (2015). *School Health Policies and Practices study*. Retrieved from http://www.

cdc.gov/healthyyouth/data/shpps/pdf/shpps-508-final_101315.pdf

van der Mars, H. (1989). Basic recording tactics. In P. W. Darst, D. Zakrajsek & V. H. Mancini (Eds.), *Analyzing physical education and sport instruction* (2nd ed., pp. 19–53). Champaign, IL: Human Kinetics.

van der Mars, H., & Harvey, S. (2010). Teaching and assessing racquet games using "Play Practice" – Part 2. *Journal of Physical Education, Recreation, and Dance, 81*(5), 35–43, 56. https://doi.org/10.1080/07303084.2010.10598478

Veal, M. L. (1988). Pupil assessment perceptions and practices of secondary teachers. *Journal of Teaching in Physical Education, 7,* 327–342. https://doi.org/10.1123/jtpe.7.4.327

Veal, M. L. (1992). The role of assessment in secondary physical education: A pedagogical view. *Journal of Physical Education, Recreation, and Dance, 63*(7), 88–92. https://doi.org/10.1080/07303084.1992.10609932

Veal, M. L. (1995). Assessment as an instructional tool. *Strategies, 8*(6), 10–15. https://doi.org/10.1080/08924562.1995.10592045

Veldhuis, M., & van den Heuvel-Panhuizen, M. (2014). Exploring the feasibility and effectiveness of assessment techniques to improve student learning in primary mathematics education. In P. Liljedahl, C. Nicol, S. Oesterle, & D. Allan (Eds.), *Proceedings of the 38th Conference of the International Group for the Psychology of Mathematics Education and the 36th Conference of the North American Chapter of the Psychology of Mathematics Education* (Vol. 5, pp. 329–336). Retrieved from http://www.pmena.org/pmenaproceedings/PMENA%2036%20PME%2038%202014%20Proceedings%20Vol%201.pdf

Weinberg, H. (1996). Authentic assessment in middle school physical education. *Teaching Secondary Physical Education, 2*(2), 20–21.

Wood, T. (2003). Assessment in physical education: The future is now! In S. Silverman & C. Ennis (Eds.), *Student learning in physical education: Applying research to enhance instruction* (2nd ed., pp. 187–203). Champaign, IL: Human Kinetics.

Appendix

SOFAST Observation Form

Date:_____ **School:** _____ **Grade:** _____ **IOA? Y / N**

# of Females _____ **Males** _____ **Class size:** ___

Baseline Session # _____ **Intervention Session #** _____ **Start time:** _____ **Obs. Length:** ____:____ **m:s**

Teacher: _____ **M / F** **SOFAST Observer:** _____

Lesson Location: In____Out ___ **Main Lesson Content:** _____

_____

Focus of Formal Assessment: _____

Minutes	*Int.*																	
0																		
	1	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
	2	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
	3	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
1																		
	4	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
	5	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
	6	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
2																		
	7	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
	8	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
	9	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
3																		
	10	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
	11	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
	12	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
4																		
	13	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
	14	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
	15	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
5																		
	16	F	I	P	K	M	O	0	C	M	S	M	Kn	F	S	G	O
	P. 1 Totals																	

P. 1 F+I Total= _____ ***# of missed intervals Page 1:*** _____