



Using systematic observation to assess teacher effectiveness promoting personally and socially responsible behavior in physical education

Paul M. Wright & Carol Irwin

To cite this article: Paul M. Wright & Carol Irwin (2018) Using systematic observation to assess teacher effectiveness promoting personally and socially responsible behavior in physical education, *Measurement in Physical Education and Exercise Science*, 22:3, 250-262, DOI: 10.1080/1091367X.2018.1429445

To link to this article: <https://doi.org/10.1080/1091367X.2018.1429445>



Published online: 05 Feb 2018.



Submit your article to this journal [↗](#)



Article views: 119



View Crossmark data [↗](#)



Using systematic observation to assess teacher effectiveness promoting personally and socially responsible behavior in physical education

Paul M. Wright^a and Carol Irwin^b

^aNorthern Illinois University, DeKalb, Illinois; ^bUniversity of Memphis, Memphis, Tennessee

ABSTRACT

National content standards in PE address responsibility; however, learning outcomes and teacher effectiveness in this area remain poorly defined. This study employed the Social and Emotional Learning framework and a teaching personal and social responsibility (TPSR) model fidelity instrument to address this gap. Our purpose was to examine the utility of the Tool for Assessing Responsibility-based Education (TARE) in assessing teacher effectiveness promoting responsibility. We conducted a comparative case study of three PE teachers in an urban public high school. Data sources included observations, interviews, and artifacts. Findings indicate that the TARE is a feasible instrument to assess PE teachers' effectiveness in promoting responsibility. Data were sufficient to develop distinct teacher profiles. While the TARE is a TPSR fidelity instrument, it is also a valid and reliable observation instrument that can be applied in the context of practice. Implications for professional development and research are discussed.

KEYWORDS

teaching personal and social responsibility model; national standards; teacher effectiveness; systematic observation

The United States (US) national standards for physical education (PE) assert that “The physically literate individual exhibits responsible personal and social behavior that respects self and others” (SHAPE, 2014). This constitutes a mandate for PE teachers to help students learn and practice responsible behavior. However, compared to other national standards, there is currently little focus on this aspect of PE. Many teachers struggle to articulate, promote, and assess learning outcomes related to responsibility (Hellison & Wright, 2011). At the same time, the academic literature indicates that the majority of researchers and policy makers are more focused on PE learning outcomes that connect to childhood obesity prevention (e.g., Amis, Wright, Dyson, Vardaman, & Ferry, 2012). Hence, the national standard related to responsibility is arguably the least well developed in terms of research, practice, and policy. While we do not suggest requiring teachers to use Hellison's (2011) Teaching Personal and Social Responsibility (TPSR) model, we do posit that research, practical strategies, and evaluation tools developed in connection with TPSR might be of great utility for PE teachers and programs struggling to address this curricular mandate in an intentional way. Unfortunately, the application of validated instruments in practice is rare. Therefore, the current study attempts to foster such linkages by demonstrating how a TPSR fidelity

instrument can be effective in guiding and assessing teacher effectiveness in promoting responsible behavior.

A discussion of teacher effectiveness in the current educational climate must be connected, at least partially, to student learning outcomes (McKenzie & Lounsbury, 2013; Rink, 2013; Ward, 2013). For decades, PE teachers and researchers have employed Bloom's taxonomy for learning domains which organizes learning into discrete categories, i.e., cognitive, psychomotor, and affective (Bloom, 1956). PE is perhaps the ideal subject matter in the school curriculum to foster the authentic integration of all three learning domains (Lund & Veal, 2013; Wright & Walsh, 2015). However, this wealth of opportunity may contribute to what some consider a muddled curriculum that addresses too many disparate and poorly-defined learning objectives (Pate & Hohn, 1994). The focus and quality of PE is also reduced in some cases by outside factors such as limited support and marginalization within the curriculum (Amis et al., 2012; Wright & Walsh, 2015) as well as the distraction and role conflict often faced by PE teachers at the secondary level who coach highly competitive inter-scholastic sports (Ressler, Richards, & Wright, 2016).

To increase coherence and focus in the PE curriculum, efforts are underway to better articulate student

learning outcomes that align with the national content standards (SHAPE, 2014). Recent discourse on synchronizing the national standards with learning outcomes and teacher effectiveness has focused on the psychomotor domain (e.g., McKenzie & Lounsbery, 2013; Rink, 2013; Ward, 2013). This reflects, at least partially, a growing desire for the field to position itself as a lever in the fight against childhood obesity (Amis et al., 2012; McKenzie & Lounsbery, 2013; Wright & Walsh, 2015). This trend, combined with a history of dismissing affective learning outcomes as difficult to measure (Hellison & Wright, 2011; Johnson, 2016; Lund & Veal, 2013), enables some teachers and researchers to view this aspect of student learning as a lower priority. However, the national standards refute this position (SHAPE, 2014). Attempts to clarify student learning outcomes and teacher effectiveness in PE must address the affective domain, including student responsibility.

Social and emotional learning

Of Bloom's (1956) three learning domains, the affective is perhaps the least clear and the least consistently defined. Affective learning objectives often address attitudes, feelings, values, behaviors, social skills, and dispositions. Hence, the range of variables as well as learning theories that could be applied in assessing these objectives can be difficult to manage. Lund and Veal (2013, p.5) succinctly state that in the context of PE, the affective domain is reflected in "exhibiting positive social behaviors (e.g., teamwork, fair play) and personal attitudes" (e.g., valuing physical activity). The current study uses the Social and Emotional Learning (SEL; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011) framework to add coherence to topics often attributed to the diffuse affective learning domain. Specifically, the SEL framework is used to operationalize personal and social responsibility in PE.

With sound theoretical and empirical support, as well as wide acceptance in the broader educational field, the SEL framework addresses student learning related to personal skills, social skills, and responsible decision-making (Durlak et al., 2011; Greenberg et al., 2003; Zhai, Raver, & Jones, 2015). While the learning process may involve attitudes, feelings, and cognitions, learning outcomes are operationalized as competencies. For example, personal skills are organized into competencies related to self-awareness (e.g., being aware of one's emotions) and self-management (e.g., regulating one's emotions). Social skills include competencies related to social awareness

(e.g., perspective taking and empathy) and relationship skills (e.g., communication and cooperation). Responsible decision-making involves making constructive and respectful choices about personal behavior and social interactions (CASEL, 2015). SEL competencies are concrete life skills that can be taught, practiced, and developed. The full range of SEL competencies can be addressed in PE as well as any other subject area (Jacobs & Wright, 2014).

Against the backdrop of SEL, the notion of personal and social responsibility becomes more clear. Student responsibilities can be framed as SEL competencies which are concrete learning outcomes within a comprehensive conceptual framework. Depending on the activity, context, and learner characteristics, appropriate and well defined learning objectives can be set using this approach. A recent publication from SHAPE America (2014) outlines grade-level outcomes to align with all national standards in grades K-12. Hence, with SEL as a framework and the current national content standards as a guide, student learning outcomes related to personal and social responsibility should neither be regarded as elusive nor impossible to define.

Teaching personal and social responsibility

While the SEL framework and content standards support teaching personal and social responsibility in PE, they do not provide concrete instructional strategies to guide teachers' practice. For this, we turn to the TPSR model (Hellison, 2011). According to Kirk's (2013) concept of model-based practice, TPSR is an instructional model because it helps to define preferred practice in terms of teaching and learner engagement. TPSR addresses a discrete set of learning outcomes that have relevance within any activity, unit, curriculum, or program plan. Hellison's (2011) responsibility goals (or levels) can be seamlessly aligned with the grade-level outcomes promoted by SHAPE America (2014). These key goals include the following: (1) respect for the rights and feelings of others; (2) self-motivation; (3) self-direction; (4) caring; and (5) transfer outside the gym (Hellison, 2011). Behaviors corresponding to these goals are often described as life skills, but can just as easily be framed as SEL competencies (Jacobs & Wright, 2014). Common examples include self-control, peaceful conflict resolution, effort, persistence, goal-setting, and leadership. The explicit focus on transfer distinguishes TPSR from other instructional models and from the framing of responsibility in the national standards (SHAPE, 2014). Transfer, described by Hellison (2011) as the ultimate goal of TPSR, involves students taking the lessons learned about responsibilities and applying them

in other contexts such as the home, classroom, or neighborhood.

A recent study validated the connection between TPSR practice and the SEL framework (Gordon, Jacobs, & Wright, 2016). This study focused on an after school physical activity program for disengaged middle school boys. Results demonstrated that the TPSR program was implemented with high fidelity to the model and that effective implementation fostered growth in all of the core SEL competencies, i.e., self-awareness, self-management, social awareness, relationship skills, and responsible decision-making. Moreover, the faculty and administration at the middle school indicated the program and the TPSR model more generally aligned with and supported the school's overall approach to promoting SEL. Although much of the TPSR literature reflects work in after school programs and alternative schools, the practical strategies and structures of the model can be used selectively by any PE teacher to address pedagogical decisions, implementation, and assessment related to student responsibility (e.g., Lund & Veal, 2013). Some structures from the typical TPSR lesson format may not translate readily into PE programming because those structures have been largely developed for use with smaller groups of students in extended day programs (Hellison, 2011). However, a recent systematic review of research conducted in PE settings indicates effective implementation of the TPSR model consistently fosters more responsible behaviors, positive attitudes and engagement among students (Pozo, Grao-Cruces, & Pérez-Ordás, 2016).

Regarding pedagogical strategies that are frequently used in the TPSR model, a set of nine were integrated into a fidelity instrument called the Tool for Assessing Responsibility-based Education (TARE; Wright & Craig, 2011). These teaching strategies are not exhaustive, but represent a continuum including common strategies that would be employed by any good teacher as well as more student-centered and empowerment based strategies that are seen less often in practice (Coulson, Irwin, & Wright, 2012; Hemphill, Templin, & Wright, 2015; Wright & Craig, 2011). The TARE strategies are: modeling respect, setting clear expectations, providing opportunities for success, fostering social interaction, assigning tasks, providing leadership opportunities, giving choices & voices in the program, sharing roles in assessment, and promoting transfer. These strategies are explained more fully, including examples of how they can be implemented, in the original publication of the instrument (Wright & Craig, 2011). The TARE has been integrated into the TPSR literature as a fidelity instrument and a training

tool (Coulson et al., 2012; Escartí, Gutiérrez, Pascual, & Wright, 2013; Escartí, Wright, Pascual, & Gutiérrez, 2015; Hellison, 2011; Hemphill et al., 2015; Wright, Jacobs, Ressler, & Jung, 2016).

In addition to supporting TPSR interventions, we propose the TARE is a measure that has broader relevance to PE. Therefore, a comparative case study design was used to demonstrate the utility of the TARE in assessing PE teachers' effectiveness promoting student responsibility. The research objectives in the current study were as follows: (1) demonstrating the feasibility of using the TARE to assess PE teachers' practice in terms of promoting personal and social responsibility; (2) using TARE data to develop individual teacher profiles that are sufficient to identify areas of strength and opportunities for growth; and (3) examining the extent to which individual teachers' TARE profiles relate to their general teaching effectiveness and to their students' behavior.

Method

Research design

This study employed a comparative case study design (Stake, 1995; Yin, 2013). Developing complete individual cases, then comparing and contrasting these cases was deemed the most effective approach to achieve the objectives of this study. Multiple methods were integrated to develop these cases. This research was conducted with Institutional Review Board approval.

Setting and participants

This study was conducted in a large urban public high school in the mid-south region of the USA. At the time of this study, the school's enrollment was 1,450. The majority of students were African-American (80.9%) followed by White (13.4%), Hispanic (2.4%), and Other (3.3%). Among the high schools in this large urban district, this Title I school had a strong reputation for academic achievement, student engagement, and athletics. The PE program was comprised of five teachers including one program chair. The PE program had adequate indoor and outdoor facilities. The school's principal was supportive of the PE program and regularly secured additional funds from the school's budget to support the program's equipment requests.

This school and three of its PE teachers were purposefully selected (Patton, 2005) to maximize variation and thereby enhance the richness of this comparative case study (Stake, 1995; Yin, 2013). The school was

selected because it presented a unique combination of several relevant factors. For example, as an urban school serving many students from underserved communities, it was likely that issues related to resources, parental involvement, student behavior and teacher effectiveness might present challenges (Cothran & Ennis, 1999; McCaughtry, Barnard, Martin, Shen, & Kulinna, 2006). At the same time, the school was relatively high achieving and had a reputation within the district of strong support for PE. Therefore, environmental factors would not necessarily restrict teachers' ability to deliver effective PE. For these reasons, the school appeared to be an appropriate real-world setting to test the feasibility (i.e., practicality and acceptability) of a measure for assessing teacher effectiveness.

Within this setting, three teachers (one White male, one African-American male, and one White female) were selected based on the following criteria: (1) being a licensed, full time teacher; (2) currently teaching PE classes; and (3) willingness to participate in the study. Of the five teachers who met these criteria, three were invited to participate because they maximized variety in terms of personal characteristics such as race, gender, years of experience, overall effectiveness, career trajectory, and approach to addressing the affective domain, including student responsibility. These selections were made in consultation with the school's principal and informed by the researchers' previous exposure to the school and its PE program.

The focus of this study was on the three teachers and included observation of their teaching, described in detail below. All observations were conducted during the same semester. For each of the three teachers, classes were selected for observation to represent the range of PE courses they were teaching at that point in time, e.g., lifetime wellness, team sports, or strength and conditioning. In the course of observing their teaching, approximately 150 of their students were involved. However, no identifiable student level data were collected. The students participating in the observed PE lessons were male and female, enrolled in grades nine through 12, and reflected the overall school demographics.

Data collection

Multiple methods were employed, including systematic observation, open field notes, formal and informal interviews, as well as document and artifact reviews. The research team was comprised of two researchers (the authors) and two research assistants. The following subsections describe data

sources, procedures, and the roles of the various research team members.

Tool for Assessing Responsibility-based Education (TARE)

The TARE is a systematic observation tool that uses time sampling to document the use of nine teaching strategies noted above. In 5-min intervals, an observer documents which of these strategies were observed. In the present study, one research assistant conducted live observations of 18 full lessons (six for each teacher) using the TARE. This research assistant was trained by the first author and demonstrated over 80% inter-rater agreement on the teaching strategies that comprise Part One of the TARE. The second and third parts of the TARE do not use time sampling, but rather are rated holistically at the end of a completed lesson. Part Two addresses general themes that Hellison (2011) has suggested convey a commitment to the promotion of student responsibility. These four themes are *Integration* of responsible roles into the activities, encouraging students to *Transfer* life skills and responsibilities to other settings, promoting *Empowerment* of students, and building a positive *Teacher-Student Relationship*. Each of these themes was rated on the following scale: Never = 1; Rarely = 2; Occasionally = 3; Frequently = 4; and Extensively = 5. Part Three rates student behavior for the group overall. Ratings are made for *Self-control*, *Participation*, *Effort*, *Self-direction*, and *Caring*. These are all rated using the following scale: Very Weak = 1; Weak = 2; Moderate = 3; Strong = 4; and Very Strong = 5. The initial publication of the TARE (Wright & Craig, 2011) reported between 88.3% and 100% inter-rater agreement on the teaching strategies that comprise Part One. Because Parts Two and Three are rated holistically for an overall lesson, the standard applied to assess their level of consistency was the percent of ratings within one point on the five-point scale. In the initial publication, these ratings were between 88.9% and 100% for items in Part Two and between 77.8% and 100% for items in Part Three.

General observations

The other research assistant was assigned to conduct general observations of the teachers and record these in the form of field notes. This assistant conducted live observations of 12 full lessons (four for each teacher) and avoided observing the same lessons as the TARE observer. This observer was not informed of the focus of the study or the TARE instrument. These steps were taken to ensure general observations were independent and unbiased. The second assistant was directed to take field notes recording descriptions of the lessons

including content covered, lesson context, instructional approach, student behaviors, teacher-student interactions, and affective climate. Each author made site visits and had meetings with the teachers and the school's principal about the study. Between the two authors, 10 such visits occurred and were documented in field notes.

Interviews

Multiple informal interviews (brief conversations during site visits and research meetings) occurred with each teacher. These informal interviews were captured in field notes and addressed topics such as the teachers' background, the school context, the PE program, or the lessons they were teaching. One formal interview was conducted with each teacher and the school's principal by one of the authors. These formal interviews lasted between 45 and 60 min, were conducted at a time and place of the participants' choosing, and were transcribed verbatim. These interviews followed a semi-structured format (Patton, 2005) with questions that addressed their background, teaching philosophy, approach to addressing the national standards, as well as their interpretation and strategies for promoting student responsibility. Examples of questions posed to the teachers include, "How would you describe your teaching philosophy?" and "What role do the national standards play in your teaching approach?". Examples of questions posed to the principal include, "How would you describe the role of physical education in your school's curriculum?" and "Can you describe any current initiatives or improvement goals in the physical education program?".

Artifacts

Artifacts were accumulated throughout the study and included information about the school available on the school and district website as well as newsletters and the popular press. Official curriculum documents from the district and school were reviewed as well as instructional materials, assessments, and rubrics developed by the teachers.

Data analysis

There were two phases of data analysis. The first phase involved the development of individual case studies. Each case study was comprised of a summary of TARE observations for the teacher integrated with all other data sources. The TARE data in each case were analyzed with descriptive statistics. The use of teaching strategies contained in Part One was represented as the percent of all observed

intervals in which a given strategy was noted. Means and standard deviations were calculated to describe the strength of themes contained in Part Two and student behaviors rated in Part Three. Qualitative data derived from general observations, field notes, interviews, and artifacts were analyzed using a combination of inductive and deductive strategies (Amis, 2005). For example, the inductive (data driven) approach was used to identify themes and patterns that emerged in each case to characterize the teacher and their overall style. These were balanced with a deductive (theory driven) approach that prompted us to look for specific units of meaning that related to student responsibility. After individual cases were developed, the second phase of analysis was cross-case comparison. In this phase, key findings from the individual cases were compared and contrasted. Cross-case comparison involved interpretation each teacher's strengths and opportunities for growth, pertinent contextual factors, and broader generalizations.

Trustworthiness

The case for trustworthiness is supported by the use of triangulation, peer debriefing, and member check (Patton, 2005). Several forms of triangulation were employed, including triangulation of data sources, methodologies, and researchers (Golafshani, 2003). This is highlighted by the different research assistants using distinct approaches to observing and documenting the teachers' practice. While the TARE data were the primary focus in this study, complimentary data sources were used to contextualize and validate those findings. In terms of peer debriefing, the first author took the lead in each phase of data analysis. After this preliminary analysis, the second author critically reviewed the summaries and interpretations. Conversations ensued to test assumptions, clarify interpretations, and identify oversights until the researchers reached consensus. Regarding member check, observations and early interpretations were discussed with the teachers during informal interviews. Finally, transcriptions of the formal interviews were shared with participants and none took issue with their accuracy.

Results

In this section, each case is presented first with a general description of each teacher followed by the specific results of their TARE observations. The three individual cases are followed by a cross-case comparison.

Mr. Brooks

At the time of this study, Mr. Brooks was a certified teacher in his third year of teaching PE. It was his first year at Main High School, but he had two years of prior teaching experience at another high school. In his previous position, he had also been a successful girls' varsity basketball coach. It was perceived by others and conveyed in his interview that his primary focus was coaching rather than teaching. This role conflict convinced the principal at Main High School that Mr. Brooks was not a good fit. By the end of the school year, Mr. Brooks was offered an opportunity to coach college basketball that he decided to pursue. He described his career goals and transition in this way:

My main goal in life was to play in the NBA and it narrowed when I couldn't play in the NBA. I wanted to do something pertaining to basketball and I think I made that kind of goal now that I'm the head coach. I had a lot of persistence with the girls' basketball, being used to winning one or two games to winning thirty games and going to state tournament. When this opportunity came about, it was something I couldn't pass, cuz I need to do this, and to get to my final goal - to be an NBA coach.

The principal at Main High School indicated that he had already decided he would not have renewed Mr. Brooks' contract because he appeared to be a sub-standard teacher.

Observations of Mr. Brooks' teaching were based in the PE portion of a year-long lifetime wellness class that addressed health education in a regular classroom-setting for one semester. The other semester addressed PE objectives in appropriate facilities, such as a gymnasium or outdoor fields. The classes included boys and girls who were mostly in their first year of high school, i.e., ninth grade. Although Mr. Brooks sometimes team taught with his colleagues and sometimes had his students in a classroom, all observations in this study occurred when he was the lone teacher in a gymnasium setting. According to field notes:

The typical day for Brooks' class would be roll call, stretches, a couple of laps around the gym, and then the students would go play their activity. After telling the students that they were finished with their laps, Coach Brooks usually said very little about anything pertaining to the day's activity. When class time was up, he would tell the students to put the equipment back on the cart it came off of and go get dressed.

Generally, Mr. Brooks' students spent about half of their time physically active, engaged in game play. He spent little time providing instruction or feedback to his students.

Little evidence of planning was found related to Mr. Brooks' instruction or assessment practices. His responses to questions on such topics were vague. For example, regarding his approach to addressing the national standards, he replied:

Definitely, those are great to follow if you're trying to reach your goals. So you try to teach your kids with all the national standards and those are the goals you wanna have your kids reach for. And those standards, the majority take them seriously you know. they want to get a certain grade that they need and a lot of them really do compete and challenge to get to those national standards so those national standards are very important to implement.

As for his approach to the affective learning domain, Mr. Brooks' focus was primarily on motivation and connected to his identity as a coach. He explained, "I'm a basketball coach, so I'm always pushing them for the best. So I want my kids, so called regular students, to be the same way as well ... I want them to have a sense of urgency as far as wellbeing and living". Regarding his approach to student behavior and responsibility, Mr. Brooks replied, "For me, I promised myself that I don't wanna have any problems with school kids because either they are gonna do it or not and I don't wanna beat myself over having any problems".

As seen in Table 1, the teaching strategies Mr. Brooks used with the greatest frequency were *Modeling Respect* (95%) and *Opportunities for Success* (72%). Other data sources confirmed that his interactions with the students were appropriate. Although he organized activities that most students could be successful in, he did not actively monitor or engage to make sure students of different ability and skill levels felt included and comfortable. The strategies of *Setting Expectations* and *Fostering Social Interaction* were observed in 47% of the intervals. This is consistent with his pattern of giving some direction and structure but then leaving students to play while he disengaged. Each of the remaining strategies were used in less than 10% of the intervals. The following

Table 1. Frequencies of observable teaching strategies from part one of the TARE.

	Teacher		
	Brooks	Gavin	Wolf
Modeling respect	95%	98%	100%
Setting expectations	47%	59%	100%
Opportunities for success	72%	86%	93%
Fostering social interaction	47%	14%	66%
Assigning tasks	0%	0%	82%
Leadership	3%	4%	0%
Giving choices and voices	9%	4%	82%
Role in assessment	0%	0%	0%
Transfer	0%	0%	9%

Table 2. Means and standard deviations for ratings of themes from part two of the TARE.

	Teacher		
	Brooks	Gavin	Wolf
Integration	2.17 (0.41)	1.83 (0.41)	4.50 (0.55)
Transfer	1.0 (0)	1.0 (0)	1.67 (1.03)
Empowerment	2.0 (0.63)	2.17 (0.75)	4.67 (0.52)
Teacher–student relationship	3.17 (0.75)	3.5 (1.05)	5.0 (0)

three strategies were never observed: *Assigning Tasks*, *Role in Assessment*, and *Transfer*.

As displayed in Table 2, across the lessons observed, Mr. Brooks was rated highest on *Teacher-Student Relationship* (3.17), followed by *Integration* (2.17), and *Empowerment* (2.0). His lowest rating was for *Transfer* (1.0) because no evidence was observed in any of the lessons. While he was not heavily engaged or enthusiastic as a teacher, he was not observed being negative or inappropriate in his interactions with students. Although levels of enthusiasm and engagement varied, Mr. Brooks' students were reasonably well behaved. As for TARE ratings of student behavior (see Table 3), *Self-control* was the highest rating (4.33), followed by *Participation* (3.83), *Effort* (3.33), *Self-direction* (3.17), and *Caring* (2.67).

Mr. Gavin

Mr. Gavin was in his second year of teaching at Main High School. He had recently completed his undergraduate degree and teaching certification and this was his first professional teaching position. In addition to teaching lifetime wellness classes, Mr. Gavin assisted coaching baseball and girls' basketball. Various data sources indicated Mr. Gavin was able to balance his coaching and teaching roles relatively well. At the time of this study, he demonstrated leadership in team teaching situations and he took on the role of cooperating teaching for a student teacher. Regarding his work with the student teacher, the principal noted, "He did a great job. He does a good job already, but when he had a student teacher he stepped it up; there's more accountability". Mr. Gavin was described by his principal as a "fairly young, but promising teacher".

Table 3. Means and standard deviations for ratings of students' behavior from part three of the TARE.

	Teacher		
	Brooks	Gavin	Wolf
Self-control	4.33 (0.52)	4.33 (0.82)	4.83 (0.41)
Participation	3.83 (0.41)	3.67 (0.52)	4.83 (0.41)
Effort	3.33 (0.52)	2.5 (0.55)	4.0 (1.10)
Self-direction	3.17 (0.41)	3.0 (0.89)	4.33 (0.52)
Caring	2.67 (0.52)	1.83 (0.75)	2.83 (1.33)

Observations of Mr. Gavin's teaching took place in the PE portion of lifetime wellness classes comprised primarily of ninth grade students as well as a team sports class available to students from all grade levels. He was often observed in a team teaching situation with students engaged in activities including bowling and step aerobics. As captured in field notes:

(Mr. Gavin) was the lead teacher in every class observed, but he received very little, if any, help from the other teachers who were supposed to be "team teaching" with him. That means that he usually taught three or even four classes largely by himself. Typically, the other three teachers would stand around and act more as classroom police rather than as teachers which was largely unnecessary with the typically good students they had in the gym. (Mr. Gavin) would lead the classes in warm-ups then divide up the students for the activities. After the groups were set up and moving forward with the activity, Mr. Gavin would walk among the students to provide help and/or encouragement to the students who needed it.

Mr. Gavin was actively engaged when teaching. He routinely provided instruction and provided feedback to his students. However, due to the large groups he was trying to organize, he spent substantial time dealing with management issues and transitions.

Mr. Gavin was able to share general plans for instructional units and assessments. He demonstrated a working knowledge of the national standards, but placed most of his attention on the psychomotor and cognitive domains. Regarding his assessment plan for a fitness unit, he explained, "The written part covers weightroom rules, proper form, muscle groups, and stretching exercises". This complimented psychomotor tests, which he summarized by stating, "I give them fitness tests on sit-ups, push-ups, the mile run, and heart rate before and after exercising". Mr. Gavin was not able to share any specific teaching strategies or assessments related to the affective domain, beyond encouraging "participation and dressing out". Still, his rapport and interactions with his students were positive.

Findings reported in Table 1 show that Mr. Gavin most frequently used *Modeling Respect* (98%) and *Opportunities for Success* (86%). This was consistent with other observations and interviews that confirmed he frequently had positive interactions with his students and emphasized participation for all. The next strategy in terms of frequency was *Setting Expectations* (59%); this drop in frequency was not surprising because his approach to student behavior was primarily reactive. The remaining six strategies were observed in less than one-half of the intervals. The following strategies were never observed: *Assigning Tasks*, *Role in Assessment*, and *Transfer*.

Regarding the themes represented in Part 2 of the TARE, Mr. Gavin received higher ratings on *Teacher-Student Relationship* (3.50) as compared to *Empowerment* (2.17) and *Integration* (1.83). Mr. Gavin did establish positive rapport with students. He was approachable and engaged students as individuals so a higher rating on *Teacher-Student Relationship* is consistent with other data sources. He was not observed promoting *Transfer* in any lessons. Table 3 displays the ratings of Mr. Gavin's students' behavior across six lessons. His students were rated highest on *Self-control* (4.33), followed by *Participation* (3.67), *Self-direction* (3.0), *Effort* (2.5), and *Caring* (1.83). These results are consistent with other observations that students were reasonably well-behaved even when they were not particularly engaged or enthused.

Ms. Wolf

At the time of this study, Ms. Wolf was a certified teacher with a master's degree and 10 years of teaching experience. She was the head girls' varsity basketball coach with a successful record including state level tournaments. At the time of this study she was applying for national board certification (which she subsequently received). She was active in the profession; not only attending, but presenting at conferences. Ms. Wolf received the Teacher of the Year award from her state professional association the year after observations were completed. She served as the Athletic Director at Main High School and the chair of the PE program. Her principal considered her extremely effective, stating, "Ms. Wolf is one of the best high school PE teachers I have ever known." While Ms. Wolf conveyed great passion for teaching and coaching, she acknowledged her long term goals included returning to school for a doctorate (which she has subsequently done) and teaching at the university level.

Ms. Wolf was observed teaching strength and conditioning classes in the school's weight room. This co-educational class was open to students of all grade levels, but most were in grades 11 and 12. The curriculum objectives in the class focused on physical fitness, proper technique in resistance training, and the development of individualized fitness plans. Students were encouraged to design and follow a fitness plan including aerobic and anaerobic exercises that would help them achieve their personal goals. Ms. Wolf provided detailed planning documents, instructional materials, rubrics, and handouts that were also available on a web page she set up for her class. Her assessments were formative (e.g., logs and reflections) and summative (e.g., final exams and fitness plans). Her approach to planning, instruction, and assessment addressed all

three learning domains. She gave students opportunities to provide feedback on the class, her teaching, as well as the content of specific assessments. Due to the nature of the class, students worked both independently and with partners. A description of the class captured in field notes stated:

The students were responsible for getting out the equipment that they needed, and if nobody else needed it then they were responsible for putting that equipment back up. The students warmed up on their own and helped each other in class with the work outs no matter if that person was their original work-out partner for the day or not. Ms. Wolf was always moving around the class to make sure the environment was safe, that the students were using correct technique, or suggesting new or alternate exercises to the students if the need arose.

Ms. Wolf was consistently engaged and enthusiastic in her approach. Her time was spent efficiently either providing instruction or facilitating student-centered activities. Little time was devoted to management or transitions.

In describing her teaching philosophy and general approach, Ms. Wolf spoke with passion and demonstrated a strong understanding of her content as well as pedagogy. She emphasized promoting lifelong physical activity and was able to speak to all learning domains and the national standards with great specificity. She frequently referred to original research, best practice, and specific texts in describing her curricular decisions. This was illustrated when she described her efforts to integrate a new fitness curriculum and text:

Last year, I followed that book's recommendations: three days a week of exercise, two days of class - the entire day of class [lecture]. I hated it, because my number one philosophy is they need to be as active as possible as much as possible. So, this past school year we had [consultants] come in and they showed us how to teach different concepts while they were active. So, I was like alright, this is more what I'm trying to look for. So this year, you would've seen less than ten-minute [lecture style] lessons, no more than three days a week. And they were never sitting. I loved it!

Ms. Wolf's approach to behavior management was proactive, with an emphasis on routines, consistency, and clear expectations. She explained, "Not dressing out every day, that's non-negotiable. And you will work. You may not give me your best, but you're gonna work - and they respond". Balancing her strong stance on expectations and routines, Ms. Wolf appreciated the importance of relationships, observing, "Yeah, because I have kids who never ever get in trouble who I can see in the hall or with another

teacher and be a totally different kid, and I totally believe it's in the relationship that I've built".

Table 1 displays Ms. Wolf's use of the teaching strategies assessed by the TARE. *Modeling Respect* and *Setting Expectations* were observed in 100% of the intervals. These findings are consistent with other observations indicating that Ms. Wolf was organized, gave clear instructions, proactively addressed behavioral expectations, and engaged respectfully with all her students. Other strategies used in the majority of intervals were *Opportunities for Success* (93%), *Assigning Tasks* (82%), *Giving Choices and Voices* (82%), and *Fostering Social Interaction* (66%). Student-centered activities that involved independent work on fitness plans, equipment management, and partners helping one another contributed to these ratings. *Transfer* was observed, but only in 9% of the intervals. When this was noted, it generally involved connecting the discussion of a life skill or attitude to another life context, e.g., giving effort in the classroom, following rules in society, and managing time in college. There was no observation of *Leadership* or *Role in Assessment* using the TARE observations although other data sources indicated she may have used these strategies at other times.

Ms. Wolf received high ratings on *Teacher-Student Relationship* (5.0), *Empowerment* (4.67), and *Integration* (4.5). These ratings are consistent with interview data and other observations. Her lowest rating was *Transfer* (1.83). This is consistent with other data sources indicating her promotion of transfer was evident but not extensive. The highest ratings for Ms. Wolf's students were for *Self-control* and *Participation*, both at 4.83. The next highest ratings were for *Self-direction* (4.33), *Effort* (4.0), and *Caring* (2.83). These findings are consistent with other observations indicating Ms. Wolf's students were engaged, often enjoying themselves, and having opportunities work independently as well as with partners.

Cross-case comparisons

All three teachers worked in the same PE program, yet represent different profiles in terms of their commitment to teaching and overall effectiveness. Mr. Brooks, for example, identified himself primarily as a coach with little commitment to teaching. While the other two teachers also coached, they appeared to have less difficulty with the role conflict this can present (Ressler et al., 2016). Mr. Gavin was the least experienced teacher, but showed some initiative within the program. He was engaged in teaching, but devoted minimal time to planning. Ms. Wolf, however, appeared to be an

excellent teacher who fully engaged with all aspects of the role and her profession.

Regarding the extent to which they addressed the affective domain and student responsibility, a similar pattern was seen. Mr. Brooks was the least engaged in teaching and accordingly had the least intentional approach regarding the affective domain. While Mr. Gavin displayed more engagement, his approach to teaching in the affective domain was minimal (e.g., promoting participation) and reactive in terms of addressing student behavior. Again, Ms. Wolf emerged as the most high performing teacher in this regard. Although promoting psychomotor development was her primary focus, all domains were addressed intentionally in her planning, instruction, and assessments. Ms. Wolf was the only teacher who appeared to view teaching responsibility as part of her role. She was aware of the TPSR model (Hellison, 2011), but had never been trained in it or made specific attempts to implement it. Her commitment to promoting responsibility was a natural part of her teaching philosophy that was also guided by the national content standards (SHAPE, 2014).

Comparing the three teachers' use of the TARE teaching strategies (see Table 1), there are similarities and differences in their profiles. All teachers made use of the first four strategies. They had consistent and high ratings on *Modeling Respect* (ranging from 95% to 100%). There was more variation in their use of *Setting Expectations* (from 47% to 100%), *Opportunities for Success* (from 72% to 93%), and *Fostering Social Interaction* (from 14% to 66%). It should be noted that on all four of these foundational strategies, Ms. Wolf used them with the most frequency. Regarding the remaining strategies, results were even more mixed. Whereas Mr. Brooks and Mr. Gavin used neither *Assigning Tasks* nor *Transfer*, Ms. Wolf did; her use of *Assigning Tasks* was extensive (82%) in comparison to *Transfer* (9%). While Mr. Brooks and Mr. Gavin made little use of *Leadership* as a strategy, 3% and 4%, respectively, Ms. Wolf was never observed using this strategy. Mr. Brooks and Mr. Gavin made limited use of *Giving Choices and Voices* (9% and 4%, respectively) compared to Ms. Wolf, who was observed using this strategy in 82% of the intervals. None of the teachers were observed using *Role in Assessment*, although interview data and artifacts indicated that Ms. Wolf did incorporate this strategy at times.

Table 2 illustrates similarities and differences between the three teachers' profiles with regard to the ratings on the themes contained in Part 2 of the TARE. All teachers received their highest ratings on *Teacher-Student Relationship* and their lowest on *Transfer*.

Consistent with the interval level data reported in Table 1, the data in Table 2 indicate that while her use of *Transfer* was relatively low, Ms. Wolf was the only teacher addressing this topic directly. Also consistent with their use of specific teaching strategies (see Table 2), theme data for Mr. Brooks and Mr. Gavin were comparable to each other and weaker than Ms. Wolf's results. Regarding the ratings of their students' behavior, *Self-control* and *Participation* were relatively high ratings for all teachers. Once again, Mr. Brooks and Mr. Gavin had distinct but comparable profiles. Student behavior ratings in Ms. Wolf's classes were higher in every category than her counterparts. *Caring* was the behavior category receiving the lowest ratings for all three teachers.

Discussion

The intent of the current study was to demonstrate the utility of the TARE in assessing PE teachers' effectiveness promoting student responsibility. Regarding the study's first objective, the TARE proved to be a feasible method for gathering data in a naturalistic school setting to assess PE teachers' practice. The second objective in the study was achieved when these data were used to develop distinct teacher profiles that were sufficient to identify areas of strength and opportunities for growth. Finally, related to the third objective, triangulation with other data sources showed that individual teachers' TARE profiles appeared to align with other indicators of their general teaching effectiveness and to relate to their students' behavior. Taken together, the findings presented here indicate the TARE has potential as an instrument for assessing teachers' effectiveness promoting personal and social responsibility in PE settings.

Rink (2013) stresses the importance of including teacher observation as part of a comprehensive approach to assessing teacher effectiveness. However, to support data-driven decision-making, data from such observations should be gathered using valid and reliable instruments (McKenzie & Lounsbury, 2013). Moreover, such instruments and procedures should be feasible (i.e., practical and acceptable) to apply in real world settings. There are few validated measures for systematically observing and documenting teaching behaviors as they relate to the affective learning domain in PE. However, related to physical fitness instruction and levels of moderate to vigorous physical activity, the System for Observing Fitness Instruction Time (SOFIT; McKenzie, 2012; McKenzie, Sallis, & Nader, 1991) has been widely applied and proven practical and acceptable for use in PE research and evaluation. Because the TARE utilizes similar procedures (McKenzie, 2012), we

anticipated it would also prove feasible as a method for assessing teacher effectiveness in PE. In the current study, an observer conducting live TARE observations did not appear to be distracting or disruptive to the ongoing PE classes. Provided with an accurate schedule and the permission of the administration and participating teachers, it was not difficult for the observer to complete multiple observations of three different teachers in a brief period of time. Although modest in scope, the current study demonstrates that the TARE presents a feasible approach to assessing PE teachers' effectiveness promoting student responsibility in naturalistic settings.

Findings presented in the current study demonstrate that the TARE can generate individualized profiles for teachers who may vary greatly in terms of experience, teaching approach, and general effectiveness. By identifying a range of specific teaching strategies that represent preferred practice (Kirk, 2013), Part One of the TARE is particularly useful in identifying areas of strength and opportunities for growth. Such data can be used to increase teachers' awareness of their typical practice and to set improvement goals. Based on results reported here and suggestions in the literature (Escartí et al., 2013, 2015; Hellison & Wright, 2011; Wright & Craig, 2011), a professional development progression might begin with making sure a teacher is consistently employing the first several strategies from Part One of the TARE (e.g., *Modeling Respect*, *Setting Expectation*, *Opportunities for Success*, and *Fostering Social Interaction*). Next steps might involve gradually increasing and/or introducing more empowerment-based strategies such as *Leadership* and *Role in Assessment*. Different examples of using the TARE as a guide for job-embedded professional development for PE teachers can be found in the literature (Coulson et al., 2012; Escartí et al., 2015; Hemphill et al., 2015).

In this study, even the teacher with the highest TARE ratings rarely promoted transfer. In the context of PE this is understandable. While most teachers would likely agree with the importance of teaching transferable life skills, few address this issue directly (Hellison, 2011). Even in TPSR programs, conversations about transfer are often short, occurring in debriefing sessions, individual conversations, or spontaneous teachable moments (Escartí et al., 2013, 2015; Hellison, 2011; Wright & Craig, 2011). While the national standards do not place a major emphasis on transfer, this aligns with the intent of the standard and can be used to differentiate higher levels of commitment to the notion of responsibility in PE. Including less common strategies, especially *Transfer*, enhances the ability of

the TARE to discriminate between different levels of intentionality and effectiveness promoting responsibility (Wright & Craig, 2011).

We are not suggesting that PE teachers be required to implement the TPSR model (Hellison, 2011). However, this instructional model provides a set of preferred practices (Kirk, 2013) that have been developed to address a set of learning outcomes (i.e., personal and social responsibility) that are explicitly identified in the national standards. Therefore, due to curricular mandates at the national level and in virtually all states (SHAPE, 2016) teachers must promote learning in this area. It stands to reason that a fidelity instrument designed to align with TPSR can be readily applied to assessing teacher effectiveness in this area. We do not propose using the TARE in a prescriptive manner or in ways that restrict teacher creativity. Rather, this instrument provides a range of strategies teachers might employ in different ways depending on their students, context, and situational insight (Hellison & Wright, 2011; Wright & Craig, 2011). Not all teachers should be expected to “teach” responsibility the same way, however, as long as it is included in their content standards, individual PE teachers should be able to articulate and reflect on what practices they do employ to support student responsibility. Toward this end, the TARE instrument may serve as a valuable resource.

In the current study, findings indicate the case study teachers’ effectiveness in promoting responsibility was an integrated part of their overall teaching effectiveness. Data from Parts One and Two of the TARE were generally aligned. Clearly, Ms. Wolf was the most effective and intentional in promoting responsibility, followed by Mr. Gavin and then Mr. Brooks. Similar statements could be made about their effectiveness in planning, delivering quality instruction, and assessing student learning. The same pattern holds true regarding their emphasis on the affective learning domain, and perhaps more importantly, to their ability to address all learning domains in a balanced way. These findings suggest that while teacher effectiveness promoting responsibility can be assessed in isolation with the TARE, such data may also inform a more general assessment of teacher effectiveness.

There is agreement that teacher effectiveness should be defined at least partially by impact on student learning (McKenzie & Lounsbury, 2013; Rink, 2013; Ward, 2013). Findings reported here indicate students exhibited more responsible behaviors at the group level in classes taught by teachers who used more responsibility-based teaching strategies. While limited in scope, this finding is consistent with previous studies indicating students tend to display higher levels of responsibility and personal-social skills in PE classes when teachers implement the TPSR model or at

least selected strategies from the TARE (Escartí et al., 2013; Escartí, Llopis-Roig, & Wright, 2016; Escartí et al., 2015; Pascual et al., 2011; Pozo et al., 2016; Wright & Burton, 2008; Wright, Li, Ding, & Pickering, 2010). This supports the rationale for using the TARE to guide and assess teacher effectiveness. Future research, including intervention studies, should examine with more specificity the relationships between specific teaching strategies and individual student outcomes.

While this study has achieved its major objectives, there are several limitations. Firstly, although the observer gathering data with the TARE was highly qualified and had demonstrated a high degree of inter-rater reliability prior to data collection, there was only one observer gathering TARE data during the study with no additional checks for reliability. Future studies of this type should include multiple observers and ongoing reliability checks. Secondly, the teachers participating in this study were observed teaching different contents with different age groups. While this reflected the reality of their schedules, it may also have unintentionally influenced their results. Therefore, our ability to make direct comparisons of their teaching styles is limited. Finally, the members of the research team did not perceive their presence as disruptive or distracting, however, aside from data triangulation, no procedures were in place to assess the extent to which teachers and their students may have behaved differently due to our visible presence (Patton, 2005).

In conclusion, demonstrating the feasibility and utility of an instrument to assess teachers’ effectiveness promoting student responsibility is an important contribution for research and practice in PE. As a valid and reliable observation tool (Escartí et al., 2013, 2015; Wright & Craig, 2011), the TARE can be used to assess, understand, and promote PE teacher effectiveness related to national curriculum standards (SHAPE, 2014) as well as broader educational initiatives related to SEL (Durlak et al., 2011; Greenberg et al., 2003; Jacobs & Wright, 2014; Zhai et al., 2015). Researchers can use this instrument to examine the relationships between various indicators of teacher effectiveness as well as a range of student learning outcomes in PE. In addition to serving as a fidelity instrument in TPSR research, the TARE has the potential to inform best practice and guide professional development for PE teachers striving to address this aspect of the national standards (SHAPE, 2014). The real world application of validated tools to support professional development and best practice is an ideal not often achieved. Making better use of validated tools to examine and promote best practice in schools is an important way to increase the effectiveness of PE in fostering a host of student learning outcomes. The current study

contributes to the literature on measurement and evaluation in PE because it provides a practical example of how validated instruments can be utilized in the context of practice.

References

- Amis, J. (2005). Interviewing for case study research. In D. L. Andrews, D. S. Mason, & M. L. Silk (Eds.), *Qualitative methods in sports studies*, (pp.104–138). Oxford, UK: Berg.
- Amis, J., Wright, P. M., Dyson, B., Vardaman, J., & Ferry, H. (2012). Implementing childhood obesity policy in a new educational environment: The cases of Mississippi and Tennessee. *American Journal of Public Health*, *102*, 1406–1413.
- Bloom, B. S. (Ed.). (1956). *Taxonomy of Educational Objectives, the classification of educational goals Handbook I: Cognitive domain*. New York, NY: McKay.
- CASEL. (2015). SEL competencies. Retrieved from <http://casel.squarespace.com/social-and-emotional-learning/core-competencies/>
- Cothran, D. J., & Ennis, C. D. (1999). Alone in a crowd: Meeting students' needs for relevance and connection in urban high school physical education. *Journal of Teaching in Physical Education*, *18*, 234–247.
- Coulson, C. L., Irwin, C. C., & Wright, P. M. (2012). Applying hellison's responsibility model in a youth residential treatment facility: A practical inquiry project. *Agora Para La Durkeducación Física Y El Deporte*, *14*, 38–54.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*, 405–432.
- Escartí, A., Gutiérrez, M., Pascual, C., & Wright, P. M. (2013). Observación de las estrategias que emplean los profesores de educación física para enseñar responsabilidad personal y social [Observation of the strategies that physical education teachers use to teach personal and social responsibility]. *Revista De Psicología Del Deporte*, *22*, 159–166.
- Escartí, A., Llopis-Roig, R., & Wright, P. M. (2016). Assessing the implementation fidelity of a school-based Teaching Personal and Social Responsibility program in physical education and other subject areas. *Journal of Teaching in Physical Education*. doi:10.1123/jtpe.2016-0200
- Escartí, A., Wright, P. M., Pascual, C., & Gutiérrez, M. (2015). Tool for Assessing Responsibility-based Education (TARE) 2.0: Instrument revisions, inter-rater reliability, and correlations between observed teaching strategies and student behaviors. *Universal Journal of Psychology*, *3*, 55–63.
- Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report*, *8*, 597–606.
- Gordon, B., Jacobs, J., & Wright, P. M. (2016). Social and emotional learning through a Teaching Personal and Social Responsibility based after-school program designed for disengaged middle-school boys. *Journal of Teaching in Physical Education*, *35*, 358–369.
- Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., & Elias, M. J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist*, *58*(6), 466.
- Hellison, D. R. (2011). *Teaching personal and social responsibility through physical activity*. Champaign, IL: Human Kinetics.
- Hellison, D. R., & Wright, P. M. (2011). *Assessment (Chapter 11) in Hellison's teaching personal and social responsibility*. Champaign, IL: Human Kinetics.
- Hemphill, M. A., Templin, T. J., & Wright, P. M. (2015). Implementation and outcomes of a responsibility-based continuing professional development protocol in physical education. *Sport, Education and Society*, *20*, 398–419.
- Jacobs, J. M., & Wright, P. M. (2014). Social and emotional learning policies and physical education. *Strategies*, *27*, 42–44.
- Johnson, T. G. (2016). Physical activity stories: Assessing the "meaning standard" in physical education. *Journal of Physical Education, Recreation & Dance*, *87*, 11–17.
- Kirk, D. (2013). Educational value and models-based practice in physical education. *Educational Philosophy and Theory*, *45*, 973–986.
- 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.
- McCaughy, N., Barnard, S., Martin, J., Shen, B., & Kulinna, P. (2006). Teachers' perspectives on the challenges of teaching physical education in urban schools: The student emotional filter. *Research Quarterly for Exercise and Sport*, *77*, 486–498.
- McKenzie, T. L. (2012). SOFIT. *System for observing fitness instruction time. Overview and training manual*. San Diego, CA: San Diego State University.
- McKenzie, T. L., & Lounsbery, M. A. F. (2013). Physical education teacher effectiveness in a public health context. *Research Quarterly for Exercise and Sport*, *84*, 419–430.
- McKenzie, T. L., Sallis, J. F., & Nader, P. R. (1991). SOFIT: System for observing fitness instruction time. *Journal of Teaching Physical Education*, *11*, 195–205.
- Pascual, C. B., Escartí, A., Guitierrez, M., Llopis, R., Marin, D., & Wright, P. M. (2011). Exploring the implementation fidelity of a program designed to promote personal and social responsibility through physical education: A comparative case study. *Research Quarterly in Exercise and Sport*, *82*, 499–511.
- Pate, R. R., & Hohn, R. C. (1994). *Health and fitness through physical education*. Champaign, IL: Human Kinetics.
- Patton, M. Q. (2005). *Qualitative research*. Oak Brook, IL: John Wiley & Sons, Ltd.
- Pozo, P., Grao-Cruces, A., & Pérez-Ordás, R. (2016). Teaching personal and social responsibility model-based programmes in physical education: A systematic review. *European Physical Education Review*, *22*, 1–20.
- Ressler, J., Richards, K. A., & Wright, P. M. (2016). The sociopolitics of sport, physical education and school health. *International Journal of Sport Policy and Politics*. doi:10.1080/19406940.2016.1202305
- Rink, J. E. (2013). Measuring teacher effectiveness in physical education. *Research Quarterly in Exercise Science*, *84*, 407–418.
- Society for Health and Physical Educators (SHAPE) America. (2014). *National standards & grade-level outcomes for K-12 physical education*. Champaign, IL: Human Kinetics.

- Society for Health and Physical Educators (SHAPE) America. (2016). *2016 Shape of the nation report: Status of physical education in the USA*. Reston, VA: Society for Health and Physical Educators (SHAPE) America.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.
- Ward, P. (2013). The role of content knowledge in conceptions of teaching effectiveness in physical education. *Research Quarterly in Physical Education, 84*, 431–440.
- Wright, P. M., & Burton, S. (2008). Examining the implementation and immediate outcomes of a personal-social responsibility model program for urban high school students. *Journal of Teaching in Physical Education, 27*, 138–154.
- Wright, P. M., & Craig, M. W. (2011). Tool for assessing responsibility-based education (TARE): Instrument development, content validity, and inter-rater reliability. *Measurement in Physical Education and Exercise Science, 15*, 204–219.
- Wright, P. M., Jacobs, J., Ressler, J., & Jung, J. (2016). Using critical pedagogy to foster transformative educational experience in a sport for development program. *Sport, Education and Society, 21*, 531–548.
- Wright, P. M., Li, W., Ding, S., & Pickering, M. (2010). Integrating a personal-social responsibility program into a lifetime wellness course for urban high school students: Assessing implementation and educational outcomes. *Sport, Education, and Society, 15*, 277–298.
- Wright, P. M., & Walsh, D. S. (2015). Subject matters of physical education. In M. F. He, B. D. Schultz, & W. H. Schubert (Eds.), *Guide to curriculum in education* (pp. 70–77). Thousand Oaks, CA: Sage.
- Yin, R. K. (2013). *Case study research: Design and methods*. Thousand Oaks, CA: Sage.
- Zhai, F., Raver, C. C., & Jones, S. M. (2015). Social and emotional learning services and child outcomes in third grade: Evidence from a cohort of Head Start participants. *Children and Youth Services Review, 56*, 42–51.