Instructional and learning outcomes in China and the USA as policy implications

By: Haiyong Ding and Ang Chen

Ding, H., & Chen, A. (2019). Instructional and learning outcomes between China and the U.S. as policy implications. *European Physical Education Review, 25*(1), 21-34.
https://doi.org/10.1177/1356336X17700165

Abstract:

There are substantial differences between the Chinese and US education systems. One difference is in the design and use of assessment due to differences in educational policies. This study described the differences in student learning outcomes, instructional approaches, and learner motivation in physical education as consequences of the policies in the two countries. Objectively measured data on skill and knowledge achievement and instructional procedures, and self-report data on student motivation were collected from a random sample of 870 students in 24 whole classes from eight Chinese middle schools and 1213 students in 39 classes from 15 US middle schools. Multivariate analyses of variance on class means revealed that Chinese students outperformed their US peers in skills, perhaps because skills were part of the high-stake tests for advancement in schooling. They were outperformed by the US students in a fitness knowledge test, perhaps because knowledge was not part of the high-stake tests. The differences in learning outcomes, instructional approaches, and motivation seem to suggest strong differentiated influences from the two countries' respective educational environments and assessment policies.

Keywords: Physical education | high-stake tests | cross-nation comparison | learning outcomes

Article:

Introduction

As early as the late 1970s, scholars from western countries became aware of the unique policy and curriculum characteristics of the educational system in modern China. Doughty (1978), for example, described China's national education doctrine of "The Three Excellences" that emphasized excellence in moral, intellectual, and physical development. He noted while the government acknowledged the need for synchronized and simultaneous development of the three, it was willing to accept an educational system that emphasizes one over the others.

During the same period, physical education scholars in the USA also began to report what they observed in physical education in China. Tien (1978) reported that physical education was given as many as 404 instructional hours in the six academic years from middle to high school, which was similar to or more than the hours given to other subject areas, such as chemistry (334 hours) or history (438). It was only surpassed by Chinese language (1146 hours) and mathematics (1170 hours). Tien also described the Chinese national standards for achievement, such as rope- or pole-climbing for 2.80 meters for boys and 2.00 meters for girls of 10–12 years of age. In 1979, Lee and Nii (1979) went to China and observed middle school physical education lessons and extra-curricular physical activities. They reported that one middle school offered 45 minute physical education lessons twice a week and there were seven physical education specialists (all male) among the 120 staff members in the school. The school also offered two physical activity sessions per week during after-class hours. During these sessions, two of the seven physical education specialists took turns to arrange physical activities for 150 students. In another middle school, similar programs (lessons and after-class sessions) were offered with five specialists (four males, one female). The reports (Lee and Nii, 1979; Tien, 1978) described the initial policy and curricular commitment the Chinese government was making to revitalize its educational system after a 10-year long closure of all schools during the Cultural Revolution (1966–1976). These reports clearly show that the Chinese education system was governed by an accountability system that was supported by compatible policies, curricula, and instructions.

Differences between the USA and China

In recent decades, physical education in both the USA and China has faced unprecedented challenges due to increased childhood obesity and physical inactivity (Fryar et al., 2012; Ji, 2007). In response, education authorities in both countries have made concerted efforts to strengthen physical education and other measures in school to provide more opportunities for students to be physically active. In 2011, the Chinese Ministry of Education (2011) published the updated learning standards for physical education to fulfill the following goals:

> Standard 1: Develop physical fitness and master knowledge, skill, and safe exercise methods, and apply the knowledge, skill, and exercise to promote health.
>
> Standard 2: Develop interest in and persistence for physical activity and exercise and regularly participate in physical activities and enjoy exercise.
>
> Standard 3: Develop mental health and display ability to collaborate and cooperate with others.
>
> Standard 4: Develop a sense of responsibility for personal and community and conduct a healthy lifestyle.
>
> Standard 5: Value and maintain a positive attitude and an open, adaptive mind toward life.

In 2013, the US Society of Health and Physical Educators (SHAPE America, 2013) specified that physical literacy is the ultimate goal of physical education. To accomplish the goal, the SHAPE published its updated national standards for physical education, which defines a physically literate person as an individual who:

> Standard 1: demonstrates competency in a variety of motor skills and movement patterns.

Standard 2: applies knowledge of concepts, principles, strategies and tactics related to movement and performance.

Standard 3: demonstrates the knowledge and skills to achieve and maintain a health-enhancing level of physical activity and fitness.

Standard 4: exhibits responsible personal and social behavior that respects self and others.

Standard 5: recognizes the value of physical activity for health, enjoyment, challenge, self-expression and/or social interaction.

Although the standards are somewhat similar in terms of encompassing knowledge, motor skill, and physical activity competencies, and affect and responsibility, there is a drastic difference in the administrative structure and implementation efforts due to known differences in school policy and curriculum between China and the USA. In China, where a national curriculum is used, physical education/activity policies and the curriculum are based on central government mandates that designate schools with obligations and legal authorities to enhance physical education/activity opportunities (Ding et al., 2014). This centralized approach has produced an environment in which changes made in the national curriculum must be followed by all institutionalized educational entities. The mandates have resulted in swift curricular changes in only a few years. In 2011, the new national physical education standards and the companion national curriculum were published to emphasize a “health-first” focus. More recently, the Chinese Prime Minister’s office (Prime Minister Office of Chinese State Council, 2016) published a directory to mandate education agencies at all levels to enforce the 2011 national standards with rigorous accountability systems for both physical education teachers and students, adding a strong emphasis on supervising the implementation of the accountability system. Fitness development goals for children at all ages are included in the physical education assessment system. Physical education assessment is now tied to graduation at all school levels. Physical education grades and actual scores in skill and fitness tests are used by schools at all levels, including college, to evaluate applicants for admission.

In the USA, school physical education/activity policy and curriculum changes are based on recommendations that do not carry federal or state legislative authority. Physical education has never been an explicit core content area in the USA. Physical education and physical activity policies usually are made at the discretion of local education agencies or school principals. Since the implementation of No Child Left Behind, 44% of schools have reduced the time allotted to physical education (Center on Education Policy, 2007, 2008). The latest Shape of the Nation report (SHAPE America and American Heart Association [AHA], 2016) showed continued reduction of physical education time and resources in US schools. According to the report, 62% (31 of 50) of states in the USA allow substituting physical education with unrelated programs, such as Junior Reserve Officer Training Corps (ROTC) programs, interscholastic sports, marching band, cheerleading, and community sports. In addition, nearly 30% of states allow schools to apply for a complete waiver from physical education requirement and 60% of states allow students to apply for an exemption from physical education.

The assessment system used in US schools is also at the discretion of local education agencies and schools. Therefore, it is as diverse as the curriculum. Instead of focusing on skill and knowledge acquisition, the most frequently used grading practices are based on attendance, effort, attitude, and dress-out, and occasionally motor skill performance. The lack of an accountability system has been documented for a long time as a major reason for the absence of learning assessment in physical education (Veal, 1988). As Matanin and Tannehill (1994) observed, most assessments were casually carried out with easy written questions and informal "spot checks." These assessment practices have led to incongruences between students' expectations for assessment and the national and state standards for learning and have misled students to developing an "easy A" expectation for physical education (Zhu, 2015).

The policy differences between the two countries were manifested in the assessment systems and the ways that assessment information is used. We argue that the policy differences create differences in the physical education curriculum and school environment for physical activity. In the Chinese system, physical education is a high-stake tested core subject. Students are tested on skills and fitness development designated in the national curriculum and evaluated by the national standards. Assessment results are part of the evaluation for advancement in schooling, including promotion to higher grades and admission to the next higher level of schools. In the USA, assessment in physical education is based on school preferences. Assessment areas vary by schools or districts. Assessment results are reported as electives in student report cards and have no impact on students' advancement to the next level of education.

Purpose of the study

School policies on curriculum and instruction determine physical education learning outcomes. These outcomes include adequate in-class physical activity time, physical intensity of the content, and knowledge and skill acquisition and improvement. In particular, the health benefits from physical education rely in large part on the curriculum and instructional structure (Chen et al., 2012a). In this research, we intended to determine differences in instructional time and learning outcomes when a similar concept-based physical education curriculum was implemented in China and the USA. Specifically, we wanted to identify outcome differences in skill, knowledge, motivation, and in-class physical activity. To accomplish this purpose, we selected learning outcome measures that were consistent with the learning goals of the respective curricula in the two countries.

Methods

Research design and setting

The study was descriptive in nature and was conducted in major metropolitan areas in both countries. The US school district served over 150,000 students and the China district served more than 300,000. Both metropolitan areas encompassed population-concentrated urban areas and large suburban areas. Both urban and suburban schools were included in the study. A randomized sampling procedure was followed in which all middle schools in the districts were stratified on enrollment (school size) and the teacher–student ratio. The sampled districts and schools were within one standard deviation of the respective national means on the two stratification variables; therefore, the samples could be considered representative of the USA and China on the two critical measures for education. To improve the ecological validity, we did not include other stratification factors (e.g. socio-economic status and test scores) in sampling due to measurement discrepancies in the two educational systems. We assumed that possible impacts from the factors were random and the randomization sampling processes would help address the impact. We employed a pre-post assessment design to gauge changes in outcome measures and collected data on curricular and instructional variables through real-time field observations. The year-long research was approved by the Institutional Review Board and Research Ethics Committee of the respective universities in the USA and China. Informed consent forms were received from participating students' parents or guardians prior to data collection. Assent forms were received from the participating students in both countries.

The school districts offered similar physical education schedules. Students in both samples had two ~50 minute lessons per week. The schools in both countries were implementing a similar concept-based curriculum that focused on teaching health-related fitness knowledge and exercises, and sport skills including basketball, badminton, soccer, and some track and field events. The curriculum also included cognitive knowledge about physical activity and health, and a portion of health education. Differences in the content existed. For example, students in the US schools were exposed to some sport skills popular in the USA that their Chinese counterparts were not, such as American football, tennis, or lacrosse. The Chinese students, on the other hand, were learning some Chinese traditional sports required in the curriculum, such as martial arts, that were not taught in the US schools. The curriculum was taught throughout the school year.

Sample and sampling

All middle schools in both school districts were included in the pool of sampling for the study. The schools were stratified on enrollment and teacher–student ratio for sampling. Schools with similar enrollment and teacher–student ratio were grouped together to form a sampling bracket (four or five schools). The final sample was drawn from the brackets to generate a random sample representing schools of all sizes and teacher–student ratios. The procedure resulted in eight Chinese schools and 15 US schools. The samples represented both large and small schools (relative to each country's standard). Once the participating schools were selected, one whole class in each of the sixth, seventh, and eighth grades was randomly selected as the data-providing unit. The final sample included 24 classes (~36 students each) from the Chinese schools and 39 classes (~31 students each) from the US schools. The whole classes were used as the unit of analysis as well. The final data set consisted of responses from 870 Chinese and 1218 US students. Boys and girls were equally represented in both countries. Table 1 reports the demographic characteristics of the student samples.

Table 1. Student demographic characteristics by country.

Country	*n*	6th Grade	7th Grade	8th Grade	Female	Male
China	870	303 / 35%	286 / 33%	281 / 32%	426 / 49%	444 / 51%
USA	1218	428 / 35%	406 / 33%	384 / 32%	628 / 51%	590 / 49%

Variables and measures

All instruments used to measure the variables in the study originated in English and were translated into Chinese. The Chinese versions were validated with satisfactory validity and reliability evidence for the Chinese participants (Ding et al., 2011, 2013).

Instruction observation. A total of 92 lessons were observed in China and 138 in the USA. Instruction was operationalized as time distributions of instructional tasks. The time distributions were used as an indicator of the extent to which students experienced what was planned for them to experience in the lesson. The time distributions were measured using a systematic observation duration-recording instrument (Rink, 2004). Student in-class activities were recorded every 15 seconds across eight categories: management; warm-up/cool-down task; instruction task (time spent listening to teacher on content-related instruction); cognitive task (e.g. completing pencil–paper assignment, measuring heart rate); fitness task (tasks exclusively declared for fitness development); skill task; game; and off-task.

Another measure used as an indicator of instruction was in-class physical activity. Learners in physical education are expected to experience and learn through developmentally appropriate physical activity at the moderate or vigorous physiological intensity levels. In-class physical activity intensity was measured using the metabolic equivalent (MET) unit. Students in the lessons that were observed wore RT-3 accelerometers that recorded the amount of three-dimensional physical activities in Vector Magnitude (VM) units. The VM units then were converted to METs based on students' sex, age, body height, and weight. Several validation studies (Hussey et al., 2009; Rawlands et al., 2004; Sun et al., 2008) have shown that data collected using RT3 accelerometers possess adequate validity and reliability in assessing children's physical activity in both physical education and free-living conditions.

Learning outcome variables. The learning outcome variables included fitness knowledge and psychomotor skills. The knowledge test was standardized and validated in both countries using the known-group method (Ding et al., 2011). The index of difficulty was between 45% and 60% and the index of discrimination was above 40% for all questions. There were 11 questions for the sixth, 10 for the seventh, and 13 for the eighth grade students. An example question for the sixth grade is: "The ability of the heart, lungs, and blood vessels to function efficiently when a person exercises the body is called…" (correct answer: cardio-respiratory fitness). An example question for the seventh grade is: "Which of the following is NOT a benefit of weight training?" (correct answer: significant increase in cardiovascular efficiency). An example question for the eighth grade is: "Alternately performing sets of exercises that train opposing muscles, without resting between sets is known as…" (correct answer: compound sets).

The AAHPERD (American Alliance for Health, Physical Education, Recreation and Dance, 1984) standardized basketball control dribble test was used to measure students' whole-body coordination skill and the Lockhart and McPherson (1949) badminton over-head clear test was used to assess the over-head striking motion. The two skills were chosen because of their general application in many forms of sports and physical activities adolescents might engage in (Chen et al., 2012a). The control dribble test was validated using a known-group approach that yielded a range of validity coefficients between .37 and .91 (AAHPERD, 1984). The badminton overhand clear test was validated using the expert-group criterion approach and the round-robin tournament rankings, which yielded a range of validity coefficients between .71 and .90

(Lockhart and McPherson, 1949). The test–retest reliability evidence ranged from .88 to .97 for the control dribble (AAHPERD, 1984) and was .90 for the badminton over-head clear (Lockhart and McPherson, 1949).

Learner motivation. Learner motivation is an important indicator of the relevance of the content and instruction due to the motivation content specificity (Chen et al., 2008) and the content characteristics (Chen et al., 2012a). In this study, student motivation was operationalized as manifested situational interest in the content, expectancy-beliefs for success in physical education, and perceived task-values in the content. These constructs were chosen due to their relevance for middle school physical education students (Chen et al., 2012b). Situational interest was measured using the 24-item Situational Interest Scale (Chen et al., 1999) and expectancy-beliefs and task-values were measured using the 13-item Expectancy-Value Inventory based on the Self- and Task-Perception Questionnaire developed by Eccles et al. (1984).

The validity of the Situational Interest Scale was established using the construct validation approach that rendered an explanatory variance of 66% and a set of internal consistency reliability coefficients ranging from .78 to .95. The Chinese version demonstrated adequate construct validity with a Root Mean Square Error of Approximation (RMSEA) of .09 with a narrow confidence interval ($CI_{(90\%)}$ = .028–.071). The internal consistency reliability coefficients ranged from .72 to .88. Validity of the English version of Self- and Task-Perception Questionnaire was deemed acceptable (Eccles et al., 1984) and the reliability was established with the internal consistency coefficients of .68 and .82 for expectancy-beliefs and task-values, respectively. Reliability of the Chinese version was established with a Chinese middle school sample. The internal consistency reliability coefficients were .85 for the expectancy-belief dimension and .69, .80, and .81 for the task-value dimensions (Ding et al., 2013).

Data collection

Trained data collectors in both countries collected the data. Knowledge and skill tests were given before and after the curriculum was taught. The knowledge test was administered in the classrooms. The skill tests were administered to individual students by the data collectors in physical education classes. Situational interest motivation was measured in the middle of the year and expectancy-value motivation was measured a week after the situational interest motivation. These measures were taken during physical education time but in classrooms. Lesson observations were conducted throughout the school year and physical activity intensity data were collected simultaneously with the observations. Students' body height and weight were measured individually in each semester and the information was used to calculate METs that reflected the physiological intensity (3–6 MET = moderate intensity, >6 = vigorous intensity).

Data reduction and analysis

The collected data were reduced into respective dimensions, total scores, and/or categories according to the specifications by the developers. Knowledge test scores were aggregated from answers to each question. A correct answer received a score of 1 and an incorrect answer 0. The correct scores were then summed and divided by the total number of questions to generate a percentage correct score for each student. Scores from the skill tests were aggregated into *T*-

scores according to the developers' instructions. Motivation scores were reduced to situational interest, expectancy-beliefs, and task-values by the developers' specifications. The original VM physical activity counts from the accelerometers were converted into METs/per minute based on students' sex, age, and body height and weight to represent the intensity.

Given the clustered structure (students were nested within classes), the unit of analysis was the means of the whole classes. All the measurement scores were aggregated at the class level. The composite class means were then computed and used in all data analyses. Descriptive statistics were calculated and *t*-test analyses and multivariate analyses of variance (MANOVAs) were conducted to determine the statistical significance of the observed differences between the two countries. Effect sizes were also calculated to determine the practical meaning of the statistical significance.

Results

Instructional time distribution

Analyzed data showed significant differences ($p < .05$) in the following lesson segments between China and US students. The Chinese students spent more time (17% of lesson time) on warm-up and cool-down tasks than the US students (9% of lesson time). They also spent more time listening to teachers' instructions (13%) than US students (7%). On the other hand, US students spent more time (23%) on skill development than Chinese students (15%); they also spent more lesson time playing sport games (25%) than the Chinese students (16%). Figure 1 displays all lesson time allocations between the two countries.

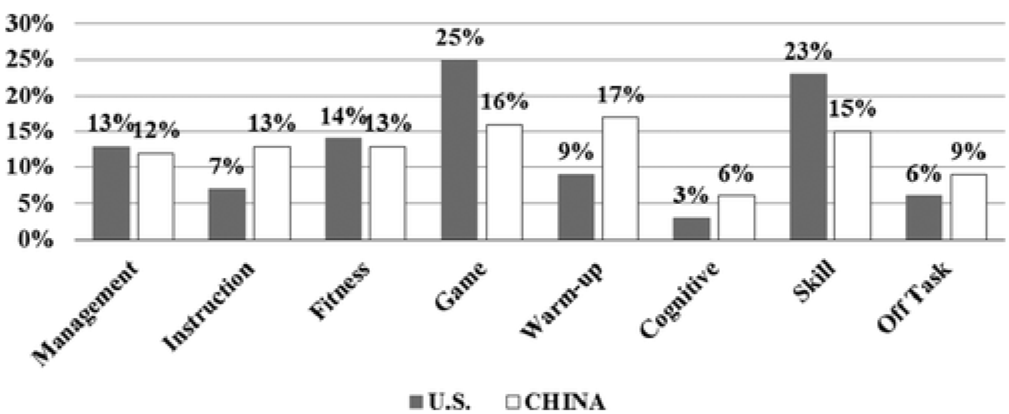

Figure 1. Average lesson time distribution comparison in lesson percentages (Chinese lessons = 92, USA lessons = 138).

Physical activity intensity

The analysis did not reveal differences between the two countries at class level. Students in both countries were similarly active. The average moderate intensity level was 3.98 METs for the US students and 4.05 METs for the Chinese students. When compared against the minute-by-minute lesson flow, as indicated in Figure 2, the physical activity intensity showed a similar pattern: warm-up was the most active segment of a lesson in both US and Chinese lessons.

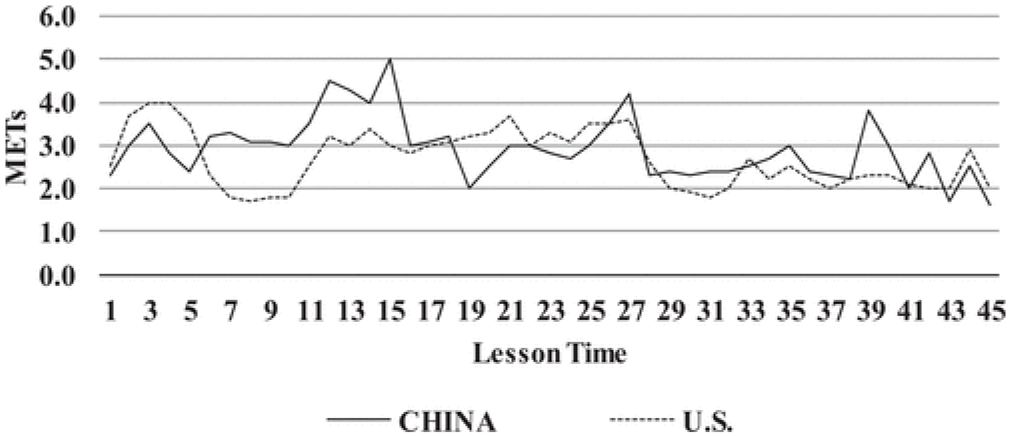

Figure 2. Average in-lesson physiological intensity comparison in metabolic equivalents (METs) (Chinese lessons = 92, US lessons = 138).

Learning outcome comparisons

Table 2 reports the descriptive statistics and analysis results for the learning outcome measures with class as the unit of analysis (24 intact classes from China and 39 from the USA). It shows that Chinese students outperformed US students in skill gains ($p < .01$). The US students demonstrated more knowledge gains ($p < .01$) by giving more correct answers to questions about health-related fitness and nutrition than the Chinese students.

Table 2. Comparison of class means/standard deviation on control dribble (*t*-scores), over-head clear (*t*-scores), and knowledge tests (% correct answer gain).

Country	*n*	Knowledge	Control dribble	Overhand clear
China	24	.39 / .15	17.40 / 5.47	25.72 / 4.71
USA	39	.43 / .17	15.97 / 4.74	21.65 / 10.99
p / *d* [a]		.007 / .25	.004 / .28	.001 / .48

[a] *t*-test *p* value and Cohen's *d*

Motivation in physical education

Students from both countries rated learning in physical education to be situationally interesting and motivating (mean for USA = 3.66, SD = 1.09, mean for China = 3.55, SD = 1.17, $p > .05$). As can be seen in Table 3, the Chinese students demonstrated weaker expectancy-beliefs for success than the Americans. However, they held stronger beliefs than the US students that physical education provided important values. Students from both countries equally appreciated the intrinsic and utility values of physical education ($p > .05$).

Table 3. Comparison of class means/standard deviation on expectancy-value motivation.

Country	*n*	Expectancy-Beliefs	Attainment value	Utility value	Intrinsic value
China	24	3.51 / .87	4.04 / .88	3.94 / 1.28	3.94 / 1.02
USA	39	4.09 / .71	3.66 / .98	3.80 / 1.01	4.07 / 1.02
p / *d* [a]		.001 / .73	.001 / .41	.063 / .12	.110 / .13

[a] Hotelling's Trace *p* value and Cohen's *d* (multivariate effect size: η^2 for Hotelling's Trace = .22).

Discussion

China may be the only country in the world where physical education is a high-stake tested core content along with mathematics, language, and science in elementary, middle, and high schools (Jones, 1999). The belief that "all that is tested is important to learn" is deeply rooted in the Chinese culture and beliefs that dominate educational policy-making processes at all levels. The centralized education system expands the influence of the culture and the beliefs profoundly. With the shift of physical education from traditional sports to knowledge and skills for health-enhancing physical activities, the centralized educational system has begun to direct school physical education toward the "health first" goal (Chinese Ministry of Education, 2011). This shift has brought significant changes in the national standards and curriculum. As Jin (2013) observed in a qualitative study, physical education teachers in China have started to embrace the "health first" goal. Another conceptual change is that the central government's policy began to encourage schools and teachers to develop "local" or "school-based" curricula. However, the extent can be limited due to the centralized system, and the effects on student learning and behavior change remain to be seen (Xu and Wong, 2011).

Differences in achievement outcomes

This study has provided initial evidence to attribute the differences in learning outcomes and instruction practices to policy differences in the USA and China. The results demonstrate the impact of the mandatory policies on pedagogy and learning outcomes in China in comparison with those in the USA. Although the instructional emphasis showed less skill/game time and more instruction time in the Chinese schools than those in the US schools, the in-class physical activity intensity levels were quite similar between the two countries. The findings indicate that physical education classes in both countries are able to provide physical activities with moderate intensity for students to receive health benefits. The findings also show that the physical education assessment systems in the two countries may have led their students to different learning outcomes.

Data from the Chinese schools appear to imply that the directive power from the centralized policy, especially the assessment policies, might have helped create an accountability system anchored on assessment. The high-stake tests for physical education seemed to have directed students' and teachers' efforts to learn and master what is to be tested. The content that was not included in the tests might have been deemed not as important. For instance, although Chinese teachers spent nearly twice as much time in verbal instruction (13% versus 7% in American lessons) and gave more cognitive tasks (7% versus 3% in American lessons), their students were still outperformed by the US students in the knowledge test. This result, speculatively, may be due to the students' realization that the fitness knowledge was not included in the high-stake test. Establishing a cause–effect relation is beyond what the data permit. It can be speculated, nevertheless, that the Chinese students placed much of their energy on developing physical fitness and improving sport skills that were tested in the high-stake tests.

For Chinese students, the high-stake tests appear to have an impact on learning sport skills. It seems reasonable to infer that the emphasis on skills from the high-stake physical education test systems (Ding et al., 2013) might have directed the Chinese students' efforts to developing sport-related psychomotor skills. The fact that they outperformed the US students on the skill tests may suggest that teaching skills for skills' sake without an accountability system, as most recreational physical education programs do in the USA (Ennis, 2010), may not be effective despite the additional time spent on skill development in physical education lessons. Speculatively, another reason could be that the Chinese teachers spent more time than the US teachers providing cognitive knowledge about the skills (see Figure 2 for the difference in instruction time where teachers provided verbal instruction to students). The additional cognitive knowledge could have made a difference in skill acquisition. Taking these potential possibilities together to contrast with the US assessment system that focused little on psychomotor skills (Matanin and Tannehill, 1994), the findings appear to imply the importance of an accountability system to hold teachers and students accountable for skill learning. It is worth pointing out that the skill tests used in the study were not the ones for physical education examination. Thus, the finding may also imply that the function of a high-stake testing system might extend beyond the skills included in the assessment system, especially considering the fact that the Chinese students spent less time on skill development in physical education lessons than the US students.

However, psychomotor skill levels were similarly low in both countries. The *T*-scores for both over-head clear and coordination tests are much below the *T*-score mean of 50. This finding should be carefully evaluated in additional studies. In both countries, physical education is deemed a necessary means to enhance health and physical activity behavior. Regardless of whether the policies are legislature-based (China) or recommendation-based (USA), they have begun to direct physical education toward a health-centered approach. The strong relation between psychomotor skills and health-related fitness development (Stodden et al., 2009) seems to be overlooked. The fact that the students from both countries scored low in skills suggests that psychomotor skills were not valued by the students. Given that both countries have changed their physical education standards recently, it remains to be seen whether the centralized policies and the national curriculum (China) or the decentralized approaches (the USA) are able to improve much needed skills for long-term health benefits.

Motivation

It is encouraging that students in both countries were almost equally motivated for learning in physical education. The finding seems to reiterate that K-12 students are motivated in general for physical education (Chen et al., 2012b). Students in both countries were equally motivated by situational interest, intrinsic value (interest and enjoyment), and utility value (usefulness). The differences observed in the study indicate that motivation may come from different sources, supporting the claim of motivation-content specificity (Chen et al., 2008). It seems that the US students relied more on expectancy-beliefs for motivation, believing they would be able to succeed in physical education. The development of this motivation can be attributable to their personal needs for competence (Wigfield and Eccles, 2002). Whether this is related to the policy and curriculum has yet to be studied.

In contrast, the Chinese students relied on the attainment value for motivation, believing physical education would enhance their health. This belief is likely to result from the policy environment in which physical education is taught and assessed. The attainment value may be instilled, reinforced, and sustained as students realize that physical education is in the high-stake test system, and therefore it must be an important subject area. Through the repeated achievement tests, Chinese students begin to see the importance of succeeding in physical education for their future education and health.

Taken together, the above findings seem to speak to the central focus of this study: there was a strong differentiated impact from policies based on legislations or government mandates (in the case of China) and policies based on recommendations by professional organizations (in the case of the USA). The findings are significant in that they show the implicit impact of different curricula and instruction as a result of the different policy-making platforms.

Conclusion

This study has provided extensive data from a large, randomized sample to demonstrate differences and commonalities in learning outcomes, instructional characteristics, and learning motivation between US and China middle school physical education. It appears that the differences may be attributable to the policy differences in the two different educational systems. The findings contribute to the literature with unique perspectives in terms of policy influences. Although physical education researchers have begun to study content choice implications for physical education, this study broadens the spectrum beyond the content with data from instructional variables, student learning variables, and student motivation variables. The findings can be meaningful for researchers, school administrators, and teachers. For researchers, the study has raised more questions than it has answered. We need to design multi-level studies to determine whether a realistic accountability system can be established to truly assist students to accomplish the goal of physical education. For school administrators, the data can help identify student needs in different policy environments in order to improve learning outcomes and motivation. For teachers, student achievement and motivation data coupled with instructional data can provide useful evidence to design innovative physical education lessons to enhance achievement and motivation.

Declaration of conflicting interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The authors received no financial support for the research, authorship, and/or publication of this article.

References

American Alliance for Health, Physical Education, Recreation and Dance [AAHPERD] (1984) AAHPERD Skills Test Manual: Basketball for Boys and Girls. Reston, VA: Author.

Center on Education Policy (2007) Choices, Changes, and Challenges: Curriculum and Instruction in the NCLB Era. Washington, DC: Author.

Center on Education Policy (2008) A Call to Restructure Restructuring: Lessons from the No Child Left Behind Act in Five States. Washington, DC: Author.

Chen, A, Darst, PW, Pangrazi, RP (1999) What constitutes situational interest? Validating a construct in physical education. Measurement in Physical Education and Exercise Science 3: 157–180.

Chen, A, Martin, R, Ennis, CD. (2008) Content specificity of expectancy beliefs and task values in elementary physical education. Research Quarterly for Exercise and Sport 79: 195–208.

Chen, A, Sun, H, Zhu, X. (2012) Influences of personal and lesson factors on caloric expenditure in physical education. Journal of Sport and Health Science 1: 49–56.

Chen, S, Chen, A, Zhu, X (2012) Are K-12 students motivated in physical education? A meta-analysis. Research Quarterly for Exercise and Sport 83: 36–48.

Chinese Ministry of Education (2011) Physical and Health Education Course Standards for Elementary and Secondary Schools. 2nd ed. Beijing: Author.

Ding, H, Li, Y, Wu, X (2014) A review of scholarly and research work in physical education in China during the first decade of the 21st century. Quest 66: 117–133.

Ding, H, Sun, H, Chen, A (2011) Gender, BMI, values, and learning in physical education: A study on Chinese middle schoolers. Learning and Individual Difference 21: 771–778.

Ding, H, Sun, H, Chen, A (2013) Impact of expectancy-value and situational interest motivation specificity on physical education outcomes. Journal of Teaching in Physical Education 32: 253–269.

Doughty, JJ (1978) The politics of education in the People's Republic of China. Theory into Practice 17: 375–382.

Eccles, JS, Adler, TF, Meece, JL (1984) In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. Personality and Social Psychology Bulletin 21: 215–225.

Ennis, CD (2010) On their own: Preparing students for a lifetime. Journal of Physical Education, Recreation, and Dance 81(5): 17–22.

Fryar, CD, Carroll, MD, Ogden, CL (2012) Prevalence of obesity among children and adolescents: United States, trends 1963–1965 through 2009–2010. Atlanta, GA: CDC National Center of Health Statistics. Available at: https://www.cdc.gov/nchs/data/hestat/obesity_child_09_10/obesity_child_09_10_htm (accessed 12 November 2015).

Hussey, J, Bennett, K, Dwyer, JO. (2009) Validation of the TR3 in the measurement of physical activity in children. Journal of Science and Medicine in Sport 12: 103–133.

Ji, C (2007) The prevalence of childhood overweight/obesity and the epidemic changes in 1985–2000 for Chinese school-age children and adolescents. Biomedical and Environmental Sciences 20(1): 1–10.

Jin, A (2013) Physical education curriculum reform in China: A perspective from physical education teachers. Physical Education and Sport Pedagogy 18: 15–27.

Jones, R (1999) Sport and physical education in school and university. In: Riordan, J, Jones, R (eds) Sport and Physical Education in China. London: International Society for Comparative Physical Education and Sport, pp.90–119.

Lee, F, Nii, C (1979) Middle school physical education in the People's Republic of China, Journal of Physical Education and Recreation 50(8): 22–65.

Lockhart, A, McPherson, FA (1949) The development of a test of badminton playing ability. Research Quarterly 20: 402–405.

Matanin, M, Tannehill, D (1994) Assessment and grading in physical education. Journal of Teaching in Physical Education 13: 395–405.

Prime Minister Office of the Chinese State Council (2016, 4) 2016 27th Directory: Strengthening school physical education for promoting student holistic, healthy development. Beijing: Chinese State Council.

Rawlands, AV, Thomas, PWM, Eston, RG. (2004) Validation of the RT3 Triaxial Accelerometer for the assessment of physical activity. Medicine & Science in Sports & Exercise 36: 518–524.

Rink, J (2004) Teaching Physical Education for Learning. 5th ed. Boston, MA: McGraw-Hill.

SHAPE America – Society of Health and Physical Educators (2013). National standards for K-12 physical education. Reston, VA.

SHAPE America and American Heart Association (2016). Shape of the nation: Status of physical education in the USA. Available at: http://www.shapeamerica.org/advocacy/son/2016/upload/Shape-of-the-Nation-2016_web.pdf (accessed 25 January 2017).

Stodden, D, Langendorger, S, Roberton, MA (2009) The association between skill competence and physical fitness in adults. Research Quarterly for Exercise and Sport 80: 229–233.

Sun, DX, Schmidt, G, Teo-Koh, SM (2008) Validation of the RT3 accelerometer for measuring physical activity of children in simulated free-living conditions. Pediatric Exercise Science 20, 181–197.

Tien, HY (1978) Body nurture: New China's physical culture and education. Theory into Practice 17: 396–400.

Veal, M (1988) Pupil assessment issues: A teacher educator's perspective. Quest 40: 151–161.

Wigfield, A, Eccles, JS (2002) The development of competence beliefs and values from childhood through adolescence. In: Wigfield, A, Eccles, JS (eds) Development of Achievement Motivation. San Diego, CA: Academic Press, pp.92–120.

Xu, Y, Wong, H (2011) School-based curriculum development in China: Policy analysis, theoretical controversy, and practical exploration. Chinese Education and Society 44(4): 44–63.

Zhu, X (2015) Student perspectives of grading in physical education. European Physical Education Review 21(4): 409–420.