Cross-Cultural Comparison of Youth Fitness Testing in China and the US: An Ecological Systems Model Approach

Xiaofen D. Keating
Department of Curriculum and Instruction, The University of Texas at Austin, xk93@utexas.edu

Rachyl Stephenson*
University of Wisconsin La Crosse, Department of Exercise and Sport Science, 1725 State Street, La Crosse, WI 54601, corresponding author, rachyl.stephenson@utexas.edu

Xiaolu Liu
Department of Curriculum and Instruction, The University of Texas at Austin, xiaolu.liu@utexas.edu

Jeff Colburn
Department of Curriculum and Instruction, The University of Texas at Austin, jeffcolburnpe@utexas.edu

Abstract

This paper takes on a different address to youth fitness testing from previous literature by applying a cross-cultural comparison of youth fitness testing in both the US and China. Utilizing a modified version of Bronfenbrenner's Ecological Model, the chronosystem (historical evolution over time), and macrosystem (social and cultural vantage point) factors, in particular, are explored within the realm of youth fitness testing. From the lens of this model changes and characteristics of youth fitness testing practice are identified. It is concluded that both fitness testing batteries were revised periodically as time passes, even though more revisions and debates occurred in the US. Although both countries made the shift from skill-related fitness to health-related fitness, social and cultural values in both countries dictated the implementation styles of youth fitness testing and the use of the current youth fitness testing results. Future theoretical research on youth fitness testing should investigate the other three factors of the model.

Keywords: school health-related fitness; social and cultural values; Bronfenbrenner's Ecological Model; physical education

There have been concerted efforts in exploring the influence of school-based physical education programs in combatting sedentarism among youth (United States Department of Health and Human Services [USDHHS], 1996, 2000). Health-related youth fitness testing, often employed within the school setting, continues to be an important predictor for health markers in children (IOM, 2012). Corbin and colleagues (2014) claimed fitness education and health-related fitness assessment are vital to physical education, the relationship between health and health-related fitness is present over all age groups, and it is necessary to teach about all aspects of health-related fitness components within a fitness education setting.

For the purpose of this paper, a cross-cultural lens is applied to compare youth fitness testing between the US and China, which are two of the largest countries in the world, representing two different educational cultures and societal systems. Cross-cultural comparisons have been made in many fields, including general education, educational psychology, and on a narrower scope in physical education. In the physical education realm, there have been a variety of studies that used the cross-cultural approach with topics such as the Theory of Planned of Behavior with physical activity (Hagger, Chatzisaranthis, Barkoukis, & Wang, 2007), exercise behaviors in college age students (Cardinal, Tuominen, & Rintala, 2004), perceived autonomy support in physical education and leisure-time physical activity (Hagger, 2005), and achievement goals in physical education (Xiang, Lee, & Solmon, 1997). All of these studies have led to advancements in our field by providing examples of global trends, or lack thereof, within the arena of physical education. By applying this to youth fitness testing, the hope is to provide significant change and advancement that has been lacking over the past 50 years as outlined by a number of researchers in the field of physical education (Corbin, 2004; Gard & Pluim, 2017; Silverman, Keating, & Phillips, 2008). With further exploration between fitness testing in these two countries, which make up almost 25% of the world's population collectively, our cross-cultural study allows us to consider different youth fitness testing perspectives, values and protocols, and analyze where policies and practices are the same, as well as where they diverge and the effectiveness of different approaches. The findings found by the present study have the potential to aide in the fight against the obesity epidemic and sedentarism in the current global context through improving student health-related fitness via fitness testing.

Theoretical Framework

Bronfenbrenner's (1979) ecological systems model was a nested system that originally consisted of four environmental levels centered around human development: microsystem, mesosystem, exosystem, and macrosystem. The model in its entirety investigates a variety of environmental factors, their complexities, and how these may impact child development or response in different situations. Bronfenbrenner outlined a microsystem as a "pattern of activities, roles, and interpersonal relations experienced by the developing person in a given setting with particular physical and material characteristics" (p. 22). Essentially, it is the smallest, most immediate environment around the child. This environment revolves around one-on-one interactions with a parent, sibling, peer, neighbor, or teacher. The mesosystem was defined as "the interrelations among two or more settings in which the developing person actively participates" (p. 25). These social interactions took place directly among the people within the environment and the focal individual (e.g. religious group, school, home, neighborhood). The exosystem signifies "one or more settings that do not involve the developing person as an active participant, but in which events occur that affect, or are affected by, what happens in the setting containing the developing person" (p. 25). These would be direct or indirect relationships, that have some sort of

trickle-down effect on the child. Examples of these settings are the local industry, school board, mass media, parents' workplace, or local government. The macrosystem represents "consistencies, in the form and content of lower-order systems that exist, or could exist, at the level of subculture or culture as a whole, along with any belief systems or ideology underlying such consistencies" (p. 26). This level within the model deals with the social and cultural impact on the child, fundamentally the ideologies of the dominant culture. From his own critical critique of the original model, Bronfenbrenner incorporated a fifth environmental level: the chronosystem. This systems is the "influence on the person's development of changes (and continuities) over time in the environments in which the person is living" (Bronfenbrenner, 1986, p. 724). The chronosystem's main focus is time, in that there are changes and constants throughout the development of a child. These changes or constants over time could pertain to personal instances such as family structure, location of the household, or parent income. However, it could also include larger contextual occurrences such as wars, economic cycles, or advancements in technology.

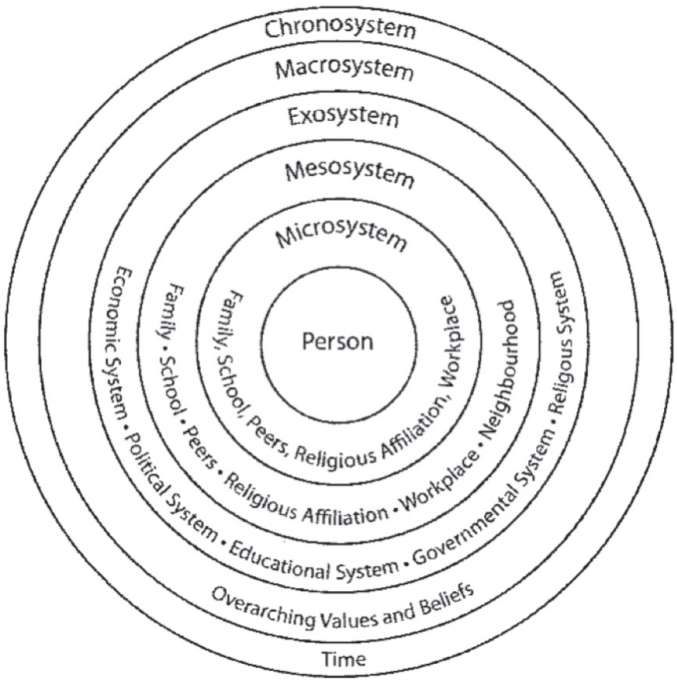

Figure 1. Bronfenbrenner's Bioecological Model

This paper aims to take an ecological model and apply it within the context of physical education to explore youth fitness testing in a cross-cultural manner between China and the US. By examining the youth fitness testing from both cultures, the hope is to have evidence of progress made, with directions on how to move forward, in order to eventually better each child's personal health and fitness growth over time. For the purpose of the paper, executing a cross-cultural perspective, only the chronosystem and macrosystem will be employed. Working from the broader environmental factors in Bronfenbrenner's model, the chronosystem, the paper will begin by examining the historical and evolutionary changes within youth fitness testing from a national and global level, beginning in the 1950s. Following is a comparison of the two cultures, and an examination of this content through Bronfenbrenner's Ecological Model. Navigating to the macrosystem level, there is a section speaking to global trends and cultural differences that need to be taken into consideration when examining youth fitness testing administration and results. There is a brief discussion of the exosystem, mesosystem, and microsystem in the future research implications section.

Chronosystem - Changes Over Time

Both China and the US have implemented systematic youth fitness testing in the school setting starting in the 1950s, but initially for different reasons (Keating, Haung, Deng, & Qu, 2003). However, in both countries, the fundamental notion behind implementing nationwide fitness testing as of late, was based on the premise that the results of these tests could support teacher use of the instruments with correct interventions, and motivation for youth to be more physically active (Fan, 1996; Morrow Jr., Zhu, Franks, Meredith, & Spain, 2009; Pangrazi, 2001). Just as the purpose of implementing fitness testing has adapted over time, so have the fitness tests themselves. Fitness testing programs in China and the US have both gone through many changes (Li, 1996; Keating, 2003; Morrow Jr. & Ede, 2009). Below is an examination of the fitness testing changes through a historical analysis in both countries. Unfortunately, the US history is much more in depth due to the lack of availability of resources and research written in English for the Chinese fitness testing history.

History and Evolution of Youth Fitness Testing in the US

In the US, the realization that American children and youth were far less fit in comparison to European youth when utilizing the Kraus-Weber assessment, initiated the start of systematic youth fitness testing (Freedson, Cureton, & Heath, 2000; Kraus & Hirschland,1953; Kraus & Hirschland,1954; Seefeldt & Vogel, 1989). President Dwight D. Eisenhower learned of the results of the initial study and thought of the impact on the military services (Zhu, Mahar, Welk, Going, & Cureton, 2011). In 1956, President's Council on Youth Fitness was created as a response to those findings. Following, the American association for Health, Physical Education, and Recreation (AHPER) organization was held many meetings on youth fitness throughout 1957-1958, and the AAHPER Fitness Test was published in 1958 (Pate, Oria, & Pillsbury, 2012). Continued attention of the link between fitness and preparedness for the military service extended into the 1960s, spilling into President-Elect John F. Kennedy's era (Morrow Jr. et al., 2009). During this time period, AAHPER was the only nationally recommended yet required test for many years, until multiple states started creating their own (Zhu et al., 2011).

The notion and implementation of health-related fitness began in the 1970s (Jackson, 2006; Pate, 1983; Pate, 1988). During this time period there was a preponderance of evidence linking fitness and physical activity to positive health outcomes and multiple committees were developed by AAHPERD (Pate et al., 2013). There were many factors that influenced this shift in fitness testing such as the publication about aerobic fitness (Cooper, 1968), and the establishment of exercise physiology, physical activity epidemiology, and measurement (Jackson, 2006). This

new commitment to fitness led AAHPERD to publish the *Health-Related Physical Fitness Test* in 1980 (Keating, 2003; Morrow Jr., Zhu, Franks, Meredith, & Spain, 2009). There was also the development of the *Youth Fitness Test*, which incorporated an awards program, administered by the retitled President's Council on Physical Fitness and Sports (PCPFS), along with a newly developed fitness report card from the Cooper Institute (Pate et al., 2013). The initial AAHPERD Health-Related Physical Fitness Test included the testing batteries of cardiorespiratory function, muscular strength and endurance, flexibility, and body composition (Zhu et al., 2011), which has basically kept the same (Keating, Smolianov, Liu, Castro-Piñero, & Smith, 2018). Throughout the 1908s, multiple revisions were done on the AAHPERD Health-Related Physical Fitness Test and Youth Fitness Test (Morrow et al., 2009; Silverman et al., 2008).

During the 1980s, and moving into the 1990s, there was a philosophical shift from norm-referenced fitness testing (compared with others) to a criterion referenced fitness testing (compared with pre-set standards). An example of a norm-referenced fitness test would be the Presidential Physical Fitness Award Program (PCPFS) where students' scores have to be above the 85th percentile on all five tests in order to achieve the award (Zhu et al., 2011). The FITNESSGRAM is an example of a criterion referenced fitness test where scores are based on a cut-off point set from an extensive literature review, originally developed in 1987 (Morrow Jr. et al., 2009; Plowman et al., 2006), with the extensive literature review for cut-off values done by Cureton (1994). There was also a call for a public health integration in youth fitness testing (Sallis & McKenzie, 1991; Simons-Morton et al., 1988). Wide scale national surveys were conducted - the *National Children and Youth Fitness Studies* (McGinnis, 1985; Ross & Pate, 1987) and the *National School Population Fitness Survey* (PCPFS, 1986). During this time period there were multiple individuals who critiqued and questioned the usage of fitness testing (Corbin, Lovejoy, Steingard, & Emerson, 1990; Freedson et al., 2000; Keating, 2003; Rowland, 1995; Seefeldt & Vogel, 1989). However, the recognition of the connection between fitness and health risk factors was as much of a motivation to refine the reliability and validity of youth fitness testing batteries (Corbin et al., 2014). The debate about the importance of relationship between physical activity and health and fitness to health factors began to dominate, reinforced the call for continuous research on refining the reliability and validity of youth fitness testing batteries (Corbin, 2007; Corbin, Pangrazi, & Welk, 1995; Silverman et al., 2008). There was an emphasis on professional organizations and governmental agencies to support policy creation as guidelines and standards for appropriate application and understanding of youth fitness testing (Pate et al., 2012).

Throughout the 90s and 2000s there were several policies, manuals, and guidelines published to support physical activity in the general population such as *Physical Activity Guidelines for Adolescent: Consensus Statement* (Sallis & Patrick, 1994), FITNESSGRAM manual with battery justification, explanation, and rationale (Morrow, Falls, & Kohl, 1994), *Complete Guide to Youth Fitness Testing* (Safrit, 1995), *Physical Activity and Health: A Report of the Surgeon General* (HHS, 1996), *Physical Activity for Children: A Statement of Guidelines* (NASPE, 1998), updating of national standards for sport and physical education (NASPE, 2004), and the physical activity guidelines for Americans announced (HHS, 2008). Over a dozen states have implemented, or are considering implementing mandatory youth fitness testing (Morrow & Ede, 2009). The National Health and Nutrition Examination Survey (NHANES) included items regarding fitness measures (Morrow Jr. et al., 2009), which allowed for longitudinal observation of fitness components and identification of the link between fitness status and health markers (Ortega, Ruiz, Castillo, & Sjostrom, 2008; Suni et al., 1998). In the past 20 years, U.S. military servicemen and women were required to adhere to a set of health-related fitness standards for enlistment and retention (IOM, 1998), along with the Department of Defense mandating annual fitness assessments (DoD, 2004). As a continuation of the findings regarding the link between physical activity and fitness in the 70s, studies during this era suggested having a physical active lifestyle has advantageous effects on a variety of fitness levels and health outcomes (Bouchard & Shephard, 1994; Pate et al., 1995; Simmons-Morton et al., 1988). Fitnessgram and the President's Challenge combined the two test programs into Fitnessgram only during 2012-2013 (Toporek, 2012). Currently, the Fitnessgram is the most widely utilized youth fitness test in the US (Castro-Pinero et al., 2010; Keating et al., 2003; Keating et al., 2018; Smolianov, Zakus, & Gallo, 2014; Welk, Going, Morrow, & Meredith, 2011). However, the percentage of state required youth fitness testing using Fitnessgram has not been significantly increased in the past 10 years (Dauenhauer et al., 2019). Due to the fact that education is administrated at the state level in the US and there are great variations in educational policies and resources across states, It is unclear to what extent that youth fitness testing practices have been in existence in the US.

In summary, nationwide mandated youth fitness testing has not been in existence in the US, even though millions of American students have taken part in fitness testing since 1950s. The widely accepted purpose for youth fitness testing is to motivate students to participate in more physical activity and adopt a lifelong physically active lifestyle. As such, physical education teachers and students are not held accountable for fitness testing results. The connection of fitness education and fitness testing is not established through policies. In addition, teachers in state with mandated fitness testing had options to choose their fitness test batteries even when the President's Challenge adopted Fitnessgram as its testing battery (Meredith & Welk, 2007). Teachers in states where fitness testing is not required can still select a different test battery (Miller et al., 2016). Moreover, testing components and items remained similar during the last 10 years (Morrorw Jr. et al., 2009; Plowan et al., 2006). Although childhood obesity has been increased dramatically in the US in recent years, the increased state mandated use of youth fitness testing is not seen, suggesting that professionals in public health and physical education have not believed that youth fitness testing could be the centerpiece in combating childhood obesity. New technology has not been widely used in field-based youth fitness test batteries in the US.

History and Evolution of Youth Fitness Testing in China

The force for integrating youth fitness testing in China was initiated to follow the model of the Soviet Union, which integrated

national fitness testing programs in schools (Li, 1996). This national fitness test, from the former Russian *Systems of Labor and Defense* in 1954 (Keating et al., 2003), was originally integrated within the military and workforce places in China. Advancements in society, along with the separation between China and Russia, lead China to establish its own youth fitness test in 1975 titled *China's National Youth Fitness Testing* and included many skill-related fitness testing items such as long jump, 50-meter run, and medicine ball throwing (Chinese Student Health Network, 2008). Through the revisions of the test battery in line with the new knowledge about health-related fitness vs. skill-related fitness, it slowly shifted to health-related fitness and renamed the China National Physical Fitness Test (CNPFT) in 2007. In 2014, over the seven-year time span from the test's origination, the test program was made more reliable and valid, and become a mandated test in schools (Liu et al., 2017).

Overall, various changes in CNPFT occurred from 1975 to 2014. These revisions included both skill and health-related fitness components within the assessment (but later placed more emphasis on the health outcomes), the number of testing items decreased through each revision, testing implementation out in the field remained relatively the same, there was an integration of technology within testing, and cut-off values decreased (became easier) over time (Chen, 2015; Liu et al., 2017). Because education is controlled at the national level in China, which is different from that in the US where each state is in charge of its own education, youth fitness testing is implemented nationwide (Keating et al., 2019; Liu et al., 2017). In addition, fitness testing results have been included in the criteria for scholarships. Graduation requirements at all educational levels also take into consideration of students' fitness testing outcomes (Keating et al., 2003). Importantly, college students are also required to take the fitness test, even though physical education is only required for freshmen and sophomores. Outside class time is used for fitness testing junior and senior students in college. However, no data are available to suggest how student fitness testing results would affect physical education instructors' performance.

Comparison Between the Histories of Youth Fitness Testing in China and the U.S.

Throughout the history of youth fitness testing, both countries made revisions in testing purpose, components, and test items. In both China and the US, the fitness testing batteries originated with skill-related fitness components (Keating et al., 2003). Based on concerns of possible negative effects of fitness testing in students, the US made changes to test items by focusing on health-related fitness components only (Morrow Jr. et al., 2009). The assessment methods also switched from having both norm- and criterion-references standards to only using a criterion-referenced testing in the 2012 (Keating et al., 2018). On the other hand, China did not convert to focus on health-related components until 2007 (Liu et al., 2017). Interestingly, neither country incorporates fitness knowledge into their testing components (Keating et al., 2003; Liu et al., 2017), even though it has been well documented that fitness knowledge plays an important role in students' fitness related behaviors (Chen, Liu, & Schaben, 2017). Fitness testing components present in both countries included body composition, cardiorespiratory fitness, muscular strength and endurance, and flexibility, but the CNPFT incorporated muscular power, whereas the US did not (Keating et al., 2018). Criterion-referenced cut-off values with age and gender differences are used to assess student fitness in both countries (Liu et al., 2017; Morrow Jr. et al., 2009). In addition, youth fitness testing has always been required in K-16 programs, even though physical education became an elective during the junior and senior years in college in China (Liu et al., 2017). Only a few states have required youth fitness testing in K-12 programs (Dauenhauer et al., 2019; Morrow Jr. et al., 2009) and college students are not required to take part in fitness testing on a regular basis in the US (Liu et al., 2017; Keating et al., 2018). The test administration is controlled by China's Department of Education whereas each state could have its own fitness testing policy (Dauenhauer et al., 2019; Keating & Silverman, 2009; Mercier, Phillips, & Silverman, 2016; Silverman et al., 2008). Furthermore, Fitnessgram is not developed by state or national educational organization(s) (Cooper Institute for Aerobics Research, 2019) while it is implemented by physical education teachers in schools if it is adopted by the state. For youth fitness testing in China, the national department of education decides on when and how the test battery should be implemented (Keating et al., 2019). To date, Fitnessgram is still only recommended and no nationally required fitness test battery exists in the US (Dauenhauer et al., 2019).

The Effect of the Chronosystem through Bronfenbrenner's Ecological Model

These changes in fitness testing represent the notion of time present in the fifth level, the chronosystem in Brofenbrenner's (1998) Ecological Model. Elder defined the principle of time and place in a historical context as, "the life course of individuals is embedded in and shaped by historical times and events they experience over their lifetime" (1998, p. 3). These changes in youth fitness testing, due to government and cultural ideological shifts, impact each child's cognitive, affective, and psychomotor development within the realm of fitness. From this perspective, the adaptations in youth fitness testing in both countries would have impact on the physical development, working all the way down to Bronfenbrenner's level 1 - the microsystem, of a child's health and growth. Within the global context, the amount and type of changes made to fitness testing between the two countries set the premise for the nested environmental side effects of a child's overall perceptions, attitudes, beliefs, and abilities in youth fitness testing.

Macrosystem - Social and Cultural Values

The above analysis of history, and comparison of youth fitness testing practices between China and the US suggest that social and cultural values greatly affect how youth fitness is assessed. During periods of war, military preparation often dictated how youth fitness testing was implemented and valued. In current day, there is not a need for war preparation, instead childhood obesity and sedentary behaviors are most impactful with youth fitness testing practices. Because of the dramatic changes in lifestyles, students are at a higher risk level for poor health more than ever before. However, except for changing fitness testing items and adjusting cut-off values for the "healthy fitness zone," field-based youth

fitness testing basically remained the same, and may have not sufficiently reflected the social and cultural values in both countries. The seemingly unstoppable increase in childhood obesity in both countries calls for more research on the empirical effects of field-based youth fitness testing practice in helping combat childhood obesity.

Concerns from a Global Context

The intent of comparing China and the US with youth fitness testing is to investigate what progress was accomplished, and what future directions may need to take place in order to support the health and fitness development of each child across the world. There is a global trend with inactivity and obesity (Skinner, Ravanbakht, Skelton, Perrin, & Armstrong, 2018), which indicates the profession of physical education and health needs to make fitness a greater priority. Macfarlane and Tomkinson (2007) warned of a potentially international public health catastrophe, if changes are not made with youth fitness testing. When deciding cut-off scores for criterion-referenced testing, age and gender had been taken into consideration in the past, but other factors such as disability and ethnicity had not (Zhu et al., 2011). The World Health Organization Multicenter Growth Reference Study Group (2006) did take into account cross-cultural differences and published international BMI standards, which is a step in the right direction. However, these reference points are norm-referenced and do not match up with the latest criterion-referenced fitness battery test style.

Moving forward, as Keating and colleagues (2018) cautioned, social and cultural factors need to be a top priority with the potential development of an international fitness testing battery. From a cross-cultural, or eventually global platform, there needs to be consideration of the cultural differences that exist such as styles and types of physical activity, and movement that represent the socialized norms within each social context. Funding and cost of equipment or test administration is another factor that will differ based on the country's economic standpoint. Fitness testing equipment, facilities, and implementation procedures could affect the accuracy of testing results and the ability to increase physical activity levels and fitness with groups of individuals who may need it most. Even within the context of youth fitness testing from a national viewpoint, there are various cultural norms. For example, in China, fitness test results can be utilized in a multitude of ways at the collegiate level such as entrance into college, applying for awards or honors, or forms of accountability (Liu et al., 2017), whereas in the US, fitness scores are not implemented for any purpose at the collegiate level. There needs to be a stronger awareness, and normalization of measuring youth fitness globally.

Discussion

In a world suffering from an obesity epidemic (Lobstein, Baur, & Uauy, 2004; Skinner et al., 2018), and a need to increase daily physical activity to combat such struggles (USDHHS 1996; 2008), youth fitness testing is an instrument that can track data in order to battle such issues (Chinese Department of Health and Fitness Education, 2016; Presidential Youth Fitness Program, 2013; President's Challenge Program, 2016). It is essential to understand how the greater societal environmental factors play into the physical development of each child. From Bronfenbrenner's Ecological Model's perspective, youth fitness testing from a historical evolution through time (chronosystem), and a cross-cultural/global social and cultural vantage point (macrosystem), youth fitness testing may have the potential to help shape the youth of today. However, given that childhood obesity has just risen over the past decades in both China and the US, it is important to examine the role of youth fitness testing in educating youth to be physically fit.

Although youth fitness testing practice has been in existence in both countries since the 1950s and a series of revisions have been made to fitness testing batteries to keep up with the changes in social and cultural values, the debates on whether youth fitness testing is needed in schools continue (Cale et al., 2007; Gard & Pluim, 2017; Seefeldt & Vogel, 1989). The long lasting common problems related to field based youth fitness testing (i.e., time consuming, forcing students to perform in front of their peers, the lack of reliability and validity of test implementations, and repetitiveness of testing items) persist in both countries since its inception in the 1950s (Chen, 2015; Garrett & Wrench, 2008; Pate, Welk, & McIver, 2013; Silverman et al., 2008). Such inability of making effective changes in youth fitness testing practice in schools, hinders the endeavor of improving students' overall health. To this end, our paper echoes the call for more research on the topic which has been made by a number of researchers (Cale et al., 2007; Ernst, Corbin, Beighle, & Pangrazi, 2006; Gard & Pluim, 2017; Silverman et al., 2008; Whitehead & Corbin, 1991). Indeed, it is time to investigate why fewer changes have been made to solve the aforementioned problems with youth fitness testing in schools.

A limitation of this paper includes the limited ability to access a wealth of information with regard to China's youth fitness testing history, due to the papers being written in Chinese. This did not allow for as in depth of a comparison between the two cultures as hoped. However, there was literature that provided enough information to display particular similarities and differences in the two countries' journeys in development and refinement of youth fitness testing.

Conclusions

Youth fitness testing on a regular basis has existed in China and the US since the 1950s. Brofenbrenner's Ecological Model is used to identify changes and characteristics of youth fitness testing practice in two of the largest countries in the world. By focusing on the chronosystem, it is concluded that both fitness testing batteries were revised periodically as time passes, even though more revisions and debates occurred in the US. Regarding the macrosystem, social and cultural values in both countries dictated the focus of the current youth fitness testing, resulting in the shift from skill-related fitness to health-related fitness.

Future Research Implications

While the focus of youth fitness testing has changed to health-related fitness, both countries linger behind swift changes in new public health challenges, caused by the dramatic increase of childhood obesity. It is disheartening that many identified common youth fitness testing problems remain unsolved for decades. There is

a need to enact new policies concerning youth fitness testing based on experimental research on the topic. Specifically, more studies are needed on how youth fitness testing can be examined through Brofenbrenner's Ecological Model within the smaller environmental contexts (exosystem, mesosystem, and microsystem). Within the exosystem level, there is a need to explore literature on how state/province and district/local policy exists or is absent pertaining to youth fitness testing. At the mesosystem level, there could be an examination of the environment that is provided for youth fitness testing within schools. This could be anywhere from the facilities and equipment, budget or funding, time allotted, class sizes, the support of administration, or the teacher involvement with youth fitness testing and the impact on the individual child. Studies are also needed concerning the characteristics of programs that have been effective in improving student fitness.

On the smallest level of the model (microsystem), there could be an exploration of factors, categorized as the demand characteristics, genetics, age, and gender that all play a vital role in the outcome of fitness testing from and individual perspective. Along with that, examining past experiences, skills required, and physical literacy necessary to execute youth fitness testing would fall under the resource characteristics of the microsystem. Another aspect of the microsystem with regard to youth fitness testing could be the force characteristics consisting of the motivation (intrinsic or extrinsic - rewards systems), ability to overcome barriers, and grit. The motivation present in the testing environments, the use of result and the programming resulting from analysis of the scores are also of concern. It would also be needed to understand what family support or culture exists with relation to youth fitness testing. With that in mind, it would be interesting to investigate in what specific cultures it is more feasible for youth fitness testing to manifest and thrive to its full potential.

New educational technology has the potential to change the landscape of education. Similar changes may also occur in youth fitness testing. As suggested by Keating and colleagues in 2018 (Keating et al., 2018), there is an urgent need to make more efforts to develop new testing items to solve the aforementioned problems associated with youth fitness testing that affect millions of students' fitness. Although the costs of using new technology may be a barrier, every effort is needed to make the use of technology a reality so that the effectiveness of youth fitness testing can be improved in both countries.

References

Bouchard, C., & Shephard, R. J. (1994). Physical activity, fitness and health: The model and key concepts: International proceedings and consensus statement. In *Physical activity, fitness and health*. Bouchard, C., Shephard, R. J., & Stephens, T.(Eds.). Champaign, IL: Human Kinetics. 11-22.

Bronfenbrenner, U. (1979). The ecology of human development: *Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (1986). Ecology of the family as a context for human development: Research perspectives. *Developmental Psychology, 22*, 723-742.

Cale, L., Harris, J., & Chen, M. H. (2007). More than 10 years after "The horse is dead...": Surely it must be time to "dismount". *Pediatric Exercise Science, 19*(2), 115-131.

Cardinal, B. J., Tuominen, K. J., & Rintala, P. (2004). Cross-Cultural comparison of American and Finnish college students' exercise behavior using transtheoretical model constructs. *Research Quarterly for Exercise and Sport, 75*(1), 92-101.

Castro-Pinero, J., Artero, E., Espana-Romero, V., Ortega, F., Sjostrom, M., Suni, J., & Ruiz, J. R. (2010). Criterion-related validity of field-based fitness tests in youth: A systematic review. *British Journal of Sports Medicine, 44*, 934-943.

Chen, H. (2015). Fitter students or flabbier standards? Thoughts and reflection on the findings of a recent Chinese fitness study. *Research Quarterly for Exercise and Sport, 86*, 213-216.

Chen, S., Liu, Y., & Schaben, J. (2017). To move more and sit less: Does physical activity/fitness knowledge matter in youth? *Journal of Teaching in Physical Education, 36*, 142-151, doi.org/10.1123/jtpe.2016-0137.

Chinese Student Health Network. (2008). Evolution and development of students' physical health assessment system. Retrieved from: http://www.csh.edu.cn/wtzx/bz/20141226/2c909e854a80abab014a8428886c0002.htm

Chinese Department of Health and Fitness Education. (2016). Physical fitness monitoring and research. Retrieved from: http://www.moe.gov.cn/s78/A17/twys_left/moe_943/

Corbin, C. B., Pangrazi, R. P., & Welk, G. J. (1995). A response to "the horse is dead; let's dismount." *Pediatric Exercise Science, 7*, 347-351.

Corbin, C. B., Lovejoy, P. Y., Steingard, P., & Emerson, R. (1990). Fitness awards: Do they accomplish their intended objectives? *American Journal of Health Promotion, 4*, 345-351.

Corbin, C. B. (2004). What every physical educator should know about teaching physical activity and fitness. *Teaching Elementary Physical Education, 15*(1), 7-9.

Corbin, C. B. (2007). Commentary to "more than 10 years after the horse is dead..." *Pediatric Exercise Science, 19*, 123-125.

Corbin, C. B., Welk, G. J., Richardson, C., Vowell, C., Lambdin, D., & Wikgran, S. (2014). Youth physical fitness: Ten key concepts. *Journal of Physical Education, Recreation & Dance, 85*(2), 24-31.

Cooper Institute for Aerobics Research. (n/d). *Cooper Institute for Aerobics Research*. Retrieved from http://www.cooperinstitute.org/

Cooper, K. H. (1968). *Aerobics*. New York, NY: Bantam Books.

Cureton, K. J. (1994). Aerobic capacity. In Morrow, J. R. Jr., Falls H. B., Kohl, H. W. III, (Eds), *The Prudential FITNESSGRAM technical reference manual*. Dallas, TX: Cooper Institute for Aerobics Research.

Dauenhauer, B., Keating, X., Stoepker, P., & Knipe, R. (2019). State physical education policy changes from 2001 to 2016. *Journal of School Health, 89*, 485-493. https://doi.org/10.1111/josh.12757

Department of Defense. (2004). *DoD physical fitness and body fat program*. Directive No. 1308.1. Washington, DC: DoD.

Elder, H. E. Jr. (1998). The life course as developmental theory. *Child Development, 69*(1), 1-12.

Ernst, M. P., Corbin, C. B., Beighle, A., & Pangrazi, R. P. (2006). Appropriate and inappropriate uses of FITNESSGRAM: A commentary. *Journal of Physical Activity & Health, 3*, S90-

S100.

Fan, L. (1996). Further improving the implementation of the National Fitness Test Program. S*chool Physical Education, 96*(2), 60-62.

Freedson, P. S., Cureton, K. J., & Heath, G. W. (2000). Status of field-based fitness testing in children and youth. *Preventive Medicine, 31*, S77-S85.

Gard, M., & Pluim, C. (2017). Why is there so little critical physical education scholarship in the United States? The case of *Fitnessgram*. *Sport, Education and Society, 22*(5), 602-617. https://doi.org/10.1080/13573322.2016.1213716

Garrett, R., & Wrench, A. (2008). Fitness testing: The pleasure and pain of it. ACHPER Hagger, M. S. (2005). Perceived autonomy support in physical education and leisure-time physical activity: A cross-cultural evaluation of the trans-contextual model. *Journal of Education Psychology, 97*(3), 376-390.

Hagger, M. S., Chatzisaranthis, N. L. D., Barkoukis, V. & Wang, J. C. K. (2007). Cross-cultural generalizability of the theory of planned behavior among young people in a physical activity context. *Exercise Psychology, 29*(1), 1-19.

Institute of Medicine. (1998). Assessing readiness in military women: *The relationship of body composition, nutrition, and health*. Washington, DC: National Academy Press.

Institute of Medicine. (2012). *Fitness measure and health outcomes in youth*. Washington, DC: The National Academies.

Jackson, A. S. (2006). The evolution and validity of health-related fitness. *Quest, 58*(1), 160-175.

Keating, X. D. (2003). The current often implemented fitness tests in physical education programs: Problems and future directions. *Quest, 55*(2), 141-160.

Keating, X. D., Huang, Y., Deng, M., & Qu, S. (2003). A comparative analysis of youth fitness test batteries between the US and the People's Republic of China. *International Sports Studies, 25*(1), 15-22.

Keating, X. D., & Silverman, S. (2009). Determinants of teacher implementation of youth fitness tests in school-based physical education programs. *Physical Education & Sport Pedagogy, 14*(2), 209-225.

Keating, X. D., Smolianov, P., Liu, X., Castro-Piñero, J., & Smith, J. (2018). Youth fitness testing practices: Global trends and new development. *The Sport Journal, 20*, 1-16.

Keating, X. D., Shangguan, R., Xiao, K., Gao, X., Sheehan, C., Wang, L., ... Wu, F. (2019). Tracking changes of Chinese pre-service teachers' aerobic fitness, body mass index, and grade point average over 4-years of college. *International Journal of Environmental Research and Public Health, 16*(6), 966. https://doi.org/10.3390/ijerph16060966.

Kraus, H., & Hirschland, R. P. (1953). Muscular fitness and health. *Journal of Physical Education, Recreation & Dance, 24*(10), 17-19.

Kraus, H., & Hirschland, R. P. (1954). Minimum muscular fitness tests in school children. *Research Quarterly, 25*, 178-188.

Li, G. (1996). The characteristics, creation, and development of the national fitness test program. *School Physical Education, 21*, 193-207.

Liu, W., He, M. Z., Wang, Y., Wang, Y., Zhou, Y., Wu, M., Tang, Z., Dai, Y., Yuan, B., Zhen, S., & Cheskin, L. J. (2017). Differences in health-related behaviors between middle school, high school, and college students in Jiangsu province, China. *Asia-Pacific Journal of Clinical Nutrition, 26*(4), 731-737.

Liu, X., Keating, X. D., & Shangguan, R. (2017). Historical analyses of fitness testing of college students in China. *ICHPER. SD Journal of Research, 9*(1), 24-32.

Lobstein, T., Baur, L., & Uauy, R. (2004). Obesity in children and young people: A crisis in public health. *Obesity Reviews, 5*(1), 84-85.

Macfarlane, D. J., & Tomkinson, G. R. (2007). Evolution and variability in fitness test performance of Asian children and adolescents. *Medical Sport Science, 50*, 143-167.

McGinnis, J. M. (1985). National Children and Youth Fitness study. *Journal of Physical Education, Recreation and Dance, 56*(1), 44-90.

Mercier, K., Phillips, S., & Silverman, S. (2016). High school physical education teachers' attitudes and use of fitness tests. *The High School Journal, 99*(2), 179-190. https://doi.org/10.1353/hsj.2016.0001

Meredith, M. D., & Welk, G. J. (Eds.). (2007). *FITNESSGRAM(r) ACTIVITYGRAM(r) test Administration Manual, The Cooper Institute* (4th ed.), Champaign, IL: Human Kinetics.

Miller, W. M., Lily, C. L., Elliot, E., Campell, H. D., Wiegand, R. L., & Bulger, S. M. (2016). Teacher perceptions of FITNESSGRAM(R) and application of results. *International Journal of Exercise Science, 9*(2), 187-204.

Morrow, J. R., Jr. & Ede, A. (2009). Statewide physical fitness testing: A big wait or a big waste? *Research Quarterly for Exercise and Sport, 80*, 696-701.

Morrow, J. R., Falls, H. B., & Kohl, H. W. (1994). *The Prudential Fitnessgram technical reference manual*. Dallas, TX: Cooper Institute for Aerobics Research.

Morrow, J. R., Zhu, W., Franks, B. D., Meredith, M. D., & Spain, C. (2009) 1958-2008: 50 years of youth fitness tests in the U.S., *Research Quarterly Exercise Sport, 80*, 1-11.

National Association for Sport and Physical Education (NASPE). (1998). *Physical activity for children: A statement of guidelines*. Reston, VA: Author.

National Association for Sport and Physical Education (NASPE). (2004). *Physical activity for children: A statement of guidelines*. Reston, VA: Author.

Oretga, F. B., Ruiz, J. R., Castillo, M. J., & Sjostrom, M. (2008). Physical fitness in childhood and adolescence: A powerful marker of health. *International Journal of Obesity, 32*(1), 1-11.

Pangrazi, R. P. (2001). *Dynamic physical education for elementary school children*. Boston: Allyn and Bacon.

Pate, R. R. (1983). A new definition of youth fitness. *Physical Sports Medicine, 11*(4), 77-83.

Pate, R. R. (1988). The evolving definition of physical fitness. *Quest, 40*(3), 174-179.

Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C.,...Wilmore, J. H. (1995). Physical activity and public health: A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *JAMA, 273*(5), 402-407.

Pate, R., Oria, M., & Pillsbury, L. (2012). *Fitness measures and*

health outcomes in youth. Washington, DC: The National Academic Press.

Pate, R. R., Welk, G. J., & McIver, K. L. (2013). Large-scale youth physical fitness testing in the United States: a 25-year retrospective review. *Pediatric Exercise Science, 25*(4), 515.

Plowman, S. A., Sterling, C. L., Corbin, C. B., Meredith, M. D., Welk, G. J., & Morrow, J. R. Jr. (2006). The history of FITNESSGRAM®. *Journal of Physical Activity & Health, 3*(Suppl. 2), S5-S20.

President's Challenge Program. (2016). President's youth fitness program (PYFP). Retrieved from: https://www.presidentschallenge.org/challenge/pyfp.shtml

President's Council on Physical Fitness and Sports. (1986). *National school population fitness survey*. Ann Arbor, MI: University of Michigan.

Presidential Youth Fitness Program. (2013). *Presidential youth fitness program physical educator resource guide*. Silver Spring, MD: National Foundation on Fitness, Sports and Nutrition.

Ross, J. G., & Pate, R. R. (1987). The National children and youth fitness study II. *Journal of Physical Education, Recreation and Dance, 58*(9), 50-96.

Rowland, T. W. (1995). The horse is dead: Let's dismount. *Pediatric Exercise Science, 7*(2), 117-120.

Safrit, M. J. (1995). *Complete guide to youth fitness testing*. Champaign, IL: Human Kinetics.

Sallis, J. F., & McKenzie, T. L. (1991). Physical education's role in public health. *Research Quarterly for Exercise and Sport, 62*, 124-137.

Sallis, J. F., & Patrick, K. (1994). Physical activity guidelines for adolescents: Consensus-statement. *Pediatric Exercise Science, 6*(4), 302-314.

Seefeldt., V., & Vogel, P. (1989). Physical fitness testing of children: A 30-year history of misguided efforts. *Pediatric Exercise Science, 1*, 295-302.

Silverman, S., Keating, X. D., & Phillips, S. R. (2008). A lasting impression: A pedagogical perspective on youth fitness testing. *Measurement in Physical Education and Exercise Science, 12*(3), 146-166.

Simons-Morton, B. G., Parcel, G. S., O'Hara, N. M., Blair, S. N., & Pate, R. R. (1988). Health-related physical fitness in childhood: Status and recommendations. *Annual Review of Public Health, 9*, 403-425.

Skinner, A. C., Ravanbakht, S. N., Skelton, J. A., Perrin, E. M., & Armstrong, S. C. (2018). Prevalence of obesity and severe obesity in US children, 1999-2016. *Pediatrics, 141*(3), e20173459. https://doi.org/10.1542/peds.2017-3459

Smolianov, P., Zakus, D., & Gallo, J. (2014). *Sport development in the United States: High performance and mass participation*. London and New York: Routledge.

Suni, J. H., Oja, P., Miilunpalo, S., I., Pasanen, M. E., Vuori, I. M., & Bos, K. (1998). Health-related fitness test battery for adults: Associations with perceived health, mobility, and back function and symptoms. *Archives of Physical Medicine and Rehabilitation, 79*(5), 559-569.

Toporek, B. (2012). *Presidential Physical Fitness test to be replaced after 2012-13*. Retrieved from http://blogs.edweek.org/edweek/schooled_in_sports/2012/09/presidential_physical_fitness_test_to_be_replaced_after_2012-13_school_year.html?cmp=SOC-SHR-FB

U.S. Department of Health and Human Services. (1996). *Physical activity and health: A report of the Surgeon General*. Atlanta, GA: Author.

U.S. Department of Health and Human Services. (2000). *Healthy People 2010: Understanding and improving health*. (2nd ed.) Washington, DC: Author.

U.S. Department of Health and Human Services. (2008). *Physical activity guidelines for Americans*. Washington, DC: Author.

Welk, G. J., Going, S. B., Morrow, J., Jr., & Meredith, M. D. (2011). Development of new criterion-referenced fitness standards in the FITNESSGRAM(r) program: Rationale and conceptual overview. *American Journal of Preventive Medicine, 41*(4S2), S63-S67.

Whitehead, J. R., & Corbin, C. B. (1991). Effects of fitness test type, teacher, and gender on exercise intrinsic motivation and physical self-worth. *Journal of School Health, 61*, 11-16. https://doi.org/10.1111/j.1746-1561.1991.tb07850.x

World Health Organization Multicentre Growth Reference Study Group. (2006). *WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development*. Geneva, Switzerland: Author.

Xiang, P., Lee, A. M., & Solmon, M. A. (1997). Achievement goals and their correlates among American and Chinese students in physical education: A cross-cultural analysis. *Journal of Cross-cultural Psychology, 28*(6), 645-660.

Zhu, W., Mahar, M. T., Welk, G. J., Going, S. B., & Cureton, K. J. (2011). Approaches for development of criterion-references standards in health-related youth fitness tests. *American Journal of Preventive Medicine, 41*(4S2), S68-S76. ■