The Official Journal of the International Society for Comparative Physical Education and Sport

PERSPECTIVE

Does 'What Works' Work? The Quest for the Holy Grail in Physical Education

Richard Peter Bailey[1]; Nadia Samsudin[1]; Hashem Salarzadeh Jenatabadi[2]; Rogayah A Razak[1]; Francis Ries[3]

[1]Faculty of Social Sciences and Liberal Arts, UCSI University, Malaysia; [2]Faculty of Science, Universiti Malaya, Malaysia; [3]Faculty of Educational Sciences, Universidad de Sevilla, Spain

Keywords:
Quality Education, Health & Well-being, Schools, Physical Activity, Sport, Evidence-based practice

Recommended Citation:
Bailey, R. P., Samsudin, N., Jenatabadi, H. S., Razak, R. A, & Ries, F. (2024). Does 'What Works' Work? The Quest for the Holy Grail in Physical Education. *International Sports Studies, 46*(2), 6-19, https://doi.org/10.69665/iss.v46i2.62

ORCID ID
Richard Peter Bailey
https://orcid.org/0000-0003-4650-6392
Nadia Samsudin
https://orcid.org/0000-0003-3292-2870
Hashem Salarzadeh Jenatabadi
https://orcid.org/0000-0002-9080-009X
Rogayah A Razak
https://orcid.org/0000-0002-2968-7745
Francis Ries
https://orcid.org/0000-0003-2569-5534

Abstract

The 'What Works' agenda within education has become popular for its focus on accountability through data, advancing the aim of developing tangible principles to improve student performance. Nevertheless, this strategy has also been criticised for seeking to define education and missing key social, emotional, and developmental issues, often narrowly. This narrative review attempts to put in perspective the use of the agenda in physical education and to combine evidence to assess its claims and shortcomings. The main conclusions suggest a dilemma in pedagogical practice between paying attention to a narrow definition of students' success in terms of physical fitness and motor skills acquisition and all the other central goals, including individual improvement, societal change, and interest in engaging in physical activity throughout life. This review takes a step towards identifying solutions to various issues in physical education by addressing evidence-based approaches and the need for interdependency with educational aims, improving the effectiveness of physical education towards students' health and well-being.

Introduction: The Rise of the 'What Works' Agenda in Education

> "Until education becomes the kind of profession that reveres evidence, we should not be surprised to find its experts dispensing unproven methods, endlessly flitting from one fad to another. The greatest victims of these fads are the very students who are most at risk." (Carnine, 2000, p. 1)

A qualitative shift in educational research, policy, and practice began in the late 20th and early 21st centuries. The rejection of the over-reliance on subjective professional judgement and an emphasis on evidence-based practices in education has now embraced what is known as the 'What Works' agenda in education research. This movement emerged from growing pressure for greater accountability and transparency in utilising public funds, in this instance, the results produced by the education sector. The 'What Works' language and style of conceptualising education emerged as educators and policy-makers focused more

on practices or strategies that seemed to lead to success. Governments, in turn, came to view this as a means of making education more systematic, reliable, and accountable to the public and policy-makers. (Lingard, 2011).

While 'What Works' was introduced for greater rigour and accountability in educational practices, this has not been without controversy. Addressing its criticisms is essential to developing a more balanced understanding of its impact, particularly in navigating the tension between measurable outcomes and the broader intrinsic goals of education.

This article aims to critically evaluate both the promises and limitations of the 'What Works' agenda, exploring its implications for policy, practice, and the holistic development of learners.

Origins and Key Players

The genesis of the 'What Works' trend in educational practices dates back to the historical advancements in social research and public policy witnessed in the 1960s and the 1970s, when the concepts of evidence-based practices started spreading in different areas such as healthcare, criminal justice, and social work (Oakley, 2000). In such fields, applying science was an attempt to determine the methods that would produce the most favourable results. As early as the 1990s, such rhetoric concerning education was developing, notably in the US and UK, where there was an emphasis on improving the education system and its output on a national scale (Slavin 2002).

The movement was profoundly shaped by prominent individuals and groups seeking to foster empirical educational research. Within the USA, the Institute of Education Sciences (IES), established in 2002 as part of the US Department of Education, became instrumental in setting benchmarks for educational research which emphasised the use of randomised controlled trials (RCTs) and other experimental methods to assess the impact of the prevalence of educational initiatives (Coalition for Evidence-Based Policy, 2002). The objective of the IES was to create a strong base for rigorous studies that would inform education policies and practices, and the organisation's emphasis on 'scientifically based research' became one of the hallmarks of the 'What Works' movement in education in North America.

Various governmental reports and programmes have fuelled the search for evidence-based education in the UK. The establishment of the 'Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre) at the University of London became influential, and it was recommended that systematic reviews of various educational research be carried out to guide policies (Oakley et al., 2005). Around the same time, the Office for Standards in Education, Children's Services, and Skills (Ofsted) implemented a new inspection framework emphasising evidence-based criteria to enhance accountability in the British educational system (Baxter and Clarke, 2013). Nowadays, for instance, the UK explicitly aims to be a world leader in pursuing evidence on 'What Works' (HM Government, 2023).

Robert Slavin (in the US) and David Hargreaves (in the UK) became well-known advocates who have played an important role in promoting the improvement agenda, the so-called 'What Works' agenda. In cooperative learning, Slavin's work stems from the extensive use of RCTs, which he considered necessary to corroborate educational interventions. For instance, treatments are scientifically tested and proven effective in medicine, and only then are they put into practice (Slavin 2008). Hargreaves also advocated evidence-based approaches. He argued that, as a matter of principle, evidence should dictate practices that cause students to learn. He stated that nothing about education can be redefined or is sacred; therefore, when making pedagogical

choices, the practices taught are grounded in sound evidence (Hargreaves, 1996).

Why 'What Works' Emerged: Accountability and Efficiency

The increasing popularity of 'What Works' in education requires explanation, and this logic can be constructed based on several factors. From the second half of the 20th Century, public and governmental agencies started to call for more efficiency and effectiveness in applying financing and spending, and education, as one of the significant cost centres, was in the spotlight. In the US and UK, dissatisfaction with students' achievements, particularly in the context of their international counterparts testing students through the Programme for International Student Assessment (PISA), created an urgency to reform the educational system (OECD, 2004). Furthermore, improving educational outcomes, providing equal opportunities for all students, and effectively appropriating public expenditure constituted a demand for evidence-based educational policy and practice (Lingard, 2011).

The inefficiencies of traditional education systems, often based on history, professional judgement, or personal experience rather than systematic evidence, have driven the rise of the 'What Works' agenda. Policy-makers began questioning whether long-standing practices were effective or necessary, particularly those intended for replication (Slavin, 2008). There has also been the application of scientific methods to social science problems in the form of RCTs and longitudinal studies, which aimed to find global best practices that could be implemented in all schools, increasing educational achievement across a wider context.

Simultaneously, the development of information technology has made it possible to manage and analyse educational data more efficiently, enabling researchers and policy-makers to monitor and evaluate educational outcomes like never before. This approach has led to the creation of extensive educational databases and longitudinal studies to identify effective teaching and policy strategies. Therefore, the intersection of policy imperatives, technological advancements, and data availability created the conditions where the 'What Works' ideology thrived and broadened the definition and the approach to achieving educational outcomes (Davies, 1999; OECD, 2004).

The Promises and Appeal of 'What Works'

According to the proponents of 'What Works', this agenda is useful in ensuring clarity and rigour in education since it focuses on identifying and documenting certain practices that can always be relied upon to improve student outcomes. Proponents argue that evidence-based practices eliminate speculation by replacing personal intuition and anecdotal evidence with data-driven decisions (Alexander, 2008). This approach aims to make educational systems more reliable by implementing proven practices to improve student outcomes.

Another attraction of the 'What Works' model is its promise of effectiveness. By identifying and implementing strategies that will help improve student outcomes, the 'What Works' agenda directly discusses how educational resources can be put to the best use. This efficiency thesis particularly appeals to makers of policy resources already in short supply, suggesting how the social rate of return on education expenditures will be maximised. Finally, the 'What Works' approach is clear and emphasises the replicability of outcomes, making the model consistent with accountability models emerging in public services. The latter becomes more important as the delivery of results becomes mandatory (Ball, 2003).

Educational Policy and the Institutionalisation of 'What Works'

The prioritisation of evidence-based practices is now fully integrated into policies and practices, particularly in the case of the Western world. Teachers are instructed to employ teaching methods that align with the prevailing norms of demonstrable and replicable success. The available resources for research and the adoption of effective practices have followed such changes in the direction of policies. A paradigm shift in the understanding of how educators should go about their craft is evident from policies that encourage or attempt to breach educators' prerogatives and impose certain approaches on them that have been validated through research (Davies, 1999; OECD, 2004).

Meanwhile, this movement has spread across many US and UK systems, where standards and accountability often drove reforms. Education researchers and policy-makers felt the influence of professions, such as medicine, which was ready to demonstrate that evidence-based approaches worked and made sense. The rationale for these changes was that education could learn from other professions since it was inefficient. Prominent policies such as the 'No Child Left Behind Act' (2001) in the US and the 'National Literacy Strategy' (1997) in the UK implemented the 'What Works' policy and established indicators that made educational success synonymous with academic achievement (Lingard, 2011; Slavin, 2008).

There is a particular risk of over-simplifying or generalising across regions and students as one modifies social problems with regard to modernist solutions, such as the 'What Works' approach. Pulling out a common schema across different practices is appealing and effective. Education systems are under unrelenting pressure to produce fair outcomes across students from different settings, even if they entail modifying variables, such as the inputs provided, where there are clear constraints, such as the school or community. Unsurprisingly, Lingard (2011) noted that this reliability has become especially appealing in a globalised world.

The central idea of the 'What Works' attitude is the introduction of the measurable results quantification paradigm in the educational sector. This shift has been transformational in curriculum methodology, teacher education, and distribution of resources within institutions. Currently, educational policy documents tend to focus on measuring learning results. The contents and methods of instruction are often framed in the context of particular competencies, such as literacy or numeracy, critical thinking, or generalisation, which are improved through effective practices (Slavin, 2008). Proponents claim that, by making such decisions based on systematic evidence, schools may minimise variability in student performance and concentrate on strategies that have demonstrated results. Such an orientation makes it possible for teachers and administrators to have a systematic approach to enhancing pupil performance to reduce differences likely to be caused by the social or academic background of the students (Davies, 1999; Gorard, 2020).

The 'What Works' agenda has also become especially attractive, as it fits into the context of accountability policies, which require schools and teachers to show progress in distinct and measurable results. Accountability has become a central concern in education, with mechanisms designed to prevent resource mismanagement and prioritise student outcomes as measures of institutional effectiveness. 'What Works' is appealing because it relies on measurable, evidence-based strategies demonstrating tangible improvements (Ball, 2003). This relationship with accountability measures is also popular with the public and governments because it can be seen as progress in education and provides the basis

for the expenditure of certain funds in implementing specific educational programmes or policies. They do this as they strongly believe in transparency and accountability in their work processes and outputs. In particular, they intend to demonstrate that the approach to schooling is as evidence-based and rigorous as medicine or engineering and that any new practice is only adopted after the most stringent tests. Focusing on scientifically based and testable approaches, the 'What Works' perspective emphasises how things work within education and how institutions improve over time. Such an emphasis on systems thinking (ST) is also gaining traction in public policymaking, including the necessity for evidence-based policy, which has been in vogue for some time. Using evidence-based methods, educational systems seek to avoid guesswork and intuition and instead focus on outcomes that can be replicated in different contexts (Coe & Kime, 2019). Traditional responses to improve education outcomes that take a 'piecemeal' approach may have some success but are unlikely to solve the 'wicked problems' that education systems face worldwide. ST offers a glimpse of a different future and may help policy-makers accomplish faster and more sustained progress in education, resulting in broad outcomes for current and future generations (Ndaruhutse et al., 2019).

In the end, the 'What Works' paradigm is designed to help teachers, administrators or policy-makers in the quest to enhance their educational outcomes by reliable means. This allows schools to take the right administrative actions that cut costs without compromising educational quality and effectiveness in terms of results. Supporters maintain that such reliance on evidence is beneficial because it enhances the credibility of the educational system in general and of educators, particularly those who, as other professionals, are ready to improve their practices for the benefit of students (Slavin, 2008). Such an argument has contributed to positioning the 'What Works' approach in mainstream education policy, with growing acceptance appealing to all wanting to improve education through research.

Detractors accept that while the focus on 'What Works' makes such an agenda more robust and consistent, it also encourages narrowing education to a technical, instrumental process that fails to consider broader educational aspirations. Educational theorists such as Biesta (2010) and Pring (2004) believe that education should not be reduced to quantifiable outputs, as this would mean it is devoid of ethical, philosophical, and developmental aspects. According to Biesta (2007), education is about making students economically effective and empowering them towards growth, active participation in a democratic system, and critical reflexivity. According to Pring (2004), educational engagement with what is measurable poses a danger to critical human beings. These critics argue that education must not only focus on 'What Works' and that an evaluation of 'what is worthwhile' is required. Hence, the model must incorporate the intrinsic values and aims of education. These critics say that because the objective of the 'What Works' perspective is to ignore these underlying themes, all other objectives are regarded as secondary and in reality.

Criticisms of 'What Works' in General Education

Critics argue that emphasis on measurable outcomes neglects the ethical, philosophical, and developmental dimensions of learning. As educational theorists observe, the issue of education is not simply a polity or policy matter, nor about knowledge transfer, but a normative and transformational process of learning to have values, character and think critically (Biesta, 2010; see also Dekker & Meeter, 2022 and Thiedig, 2023). According to this view, education should ask 'What Works' and 'what is worthwhile'. From this perspective, education should develop well-

rounded people capable of critical, ethical, and civic commitment rather than just being knowledgeable about performing a set task. Biesta (2007) argues that the obsession with worded outcomes of education is detrimental as it encourages a technocratic class of teachers and students, where the measure of success is only determined by how well the task is achieved or replicated and not what impact it has on the person completing the task. According to Biesta (2010), learning is not reduced to obtaining a set of competencies or fulfilling an academic goal; instead, it is a learning process that allows the individual to form autonomous thinking and ethical judgement, especially in a world that is becoming more complex and interdependent. Similarly, Pring (2004) maintains that educational goals should not only be reduced to developing skills. These are merely instruments of doing whose purposes are much broader in nurturing socio-political commitment and living a personally responsible and moral life.

According to these critics, an over-emphasis of education on 'What Works' endangers achieving such aims and purposes, which are value-oriented and necessary for education as an institution, which may make students less able to make meaningful contributions to their communities. However, this approach is not without its criticism. One is that teachers employ universal teaching strategies. These constraints are seen as detrimental to teachers' independence and agency. It is argued that this standardisation practice is a disservice to the conceptualisation of teachers as professionals and creative individuals who can drive the needs of a diverse classroom and create a constructive yet positive sense of learning. Teachers should not only be regarded as machines who come to class to teach students concepts and go home but also as leaders who guide and shape a child's development (Hargreaves, 1996).

Critics also point out that an excessive focus on standardisation may have an adverse effect in which students are viewed instrumentally as end products that can be optimised rather than as people who need nurturing as they grow. The 'What Works' strategy risks depersonalising education by overlooking its relational and human aspects. In this model, learners risk becoming passive recipients of skills and knowledge rather than active participants in their education. This viewpoint can transform education into a business in which the end product or output the students seek must be in terms of achievements instead of the student's potential and identities (Biesta, 2007). They argue that this model ignores the need to develop natural, internal, or intrinsic motivation, explore or be curious, and even the emotional perspective needed to nurture a passion for education and self-empowerment.

A further criticism is that the 'What Works' agenda emphasises efficiency and outcomes and less on learning processes. Being concerned only with what can be measured, this approach may omit essential elements and intricacies of educational advancement. For instance, Pring (2004) posits that it is not feasible to regard educational processes as mere input-output models because they comprise a complicated web of sometimes undetermined relationships among teachers, students, and curriculum content. The 'What Works' model's reductionism creates a situation whereby the predominant understanding of learning is overly narrowed to acquiring some skills or information, neglecting education's broader cognitive, moral, and social purposes. For example, although empathy, resilience, and ethical understanding are important for a holistic education, they are difficult to measure. As such, potential students' critics worry that the 'What Works' strategy may not create well-rounded individuals because it ignores these more challenging aspects to describe.

Equity concerns have also been raised because of the nature of the data sources

used, which are largely quantitative and standardised. It has been suggested that the ‘What Works’ paradigm may further entrench inequalities since it largely prefers circumstances and results readily measured and compared across various other contexts (Lingard 2011). As an illustration, expect schools with limited resources to seek ways of achieving standardised tests, which will be detrimental in the context of target-based performance assessment. This may also adversely affect students from the periphery, who may require more focused and contextual interventions than generic approaches. However, an emphasis on a standard set of measures has been seen to prevent the focus from the essential problems confronting various types of schools and communities. Thus, the ‘What Works’ approach can be expected to worsen educational disparities instead of resolving them (Alexander, 2008). Such proponents believe that if ‘What Works’ principles are applied uncritically, this could lead to structural inequalities in education, in which social mobility and social justice will be compromised.

Finally, critics fault the ‘What Works’ approach for its reductionism, which oversimplifies the complexities of education.". While critics point out that these quantitative data are useful, it is worth noting that an overemphasis on such quantifiable things makes people mislay on other important aspects that are difficult to assess using traditional metrics. Apart from mere information, education also includes attitudes, belief systems, and social skills that are less likely to be verifiable through numbers (Pring, 2004). In its pursuit of effectiveness, the ‘What Works’ policy risks prioritising measurable outcomes over the deeper, transformative aspects of education. This concern resonates with Hargreaves’s (1996) view that it is necessary to pay attention to the ‘whatever is taken for granted in learning settings’ aspect or ‘hidden curriculum’- the unspoken and unwritten information boundaries that pupils manage to acquire through their socialisation, emotional life and culture. This is a concern when these may be deemed soft, core to authentic learning, and should not be traded off for complex values.

In a nutshell, however, what some laud as strength – the focus on hardcore measures and assessable results to gauge the success of any programme – has its detractors as well, who could be called critics of the ‘What Works’ strategy. The contention is that such a single-minded approach will promote a reductionist view of education as an exercise in some mechanistic exchanges or management completely divorced from human emotions. They argue that such pressures lead to a perversion of the “What Works” ideology, which is meant to perfect the educational learning experience, as it is critical to furnish students with the skills and capabilities for viable citizenship. They argue for an alternative model that respects qualitative perspectives and the independence of classroom practitioners and out-of-school children and, therefore, practices education as an evolving enterprise that helps people aspire to achieve much greater socio-personal goals.

The Influence of ‘What Works’ on Physical Education

Regarding physical education (PE), good examples of the ‘What Works’ movement are systematic reviews and meta-analyses, which contribute to establishing a ‘What Works’ approach. As for the introductory course of PE, the metrics-driven approach of measuring fitness, motor, and physical competence aligns with international accountability policy in the education sector. However, despite the benefits of systematic reviews and similar processes in structuring PE as a field, there has been a notable absence of debate around the 'What Works' ideology. The policy area does not appear to involve pioneer institutions and enduring controversies regarding evidence-based arguments and policies. The lack of engagement has led to PE being moulded

by external forces, whereby what is important seemed to be dictated through systematic reviews. Consequently, there is always an argument of conflict with the deployment and use of PE development standards and policies and enhancing the core purpose and primary objectives of PE, such as increased social interaction, personal development, and an overall appreciation for physical activity (PA) (Penney, 2013; Macdonald, 2013).

The advent of the 'What Works' agenda has critically influenced the focus on policies in PE, particularly stressing the importance of fitness, motor skills, and general physical competence. This is consistent with an approach in which achievement and measurements fit into the current intensified concerns with accountability and evidence-based practices. This trend is consistent with growing pressures in the educational sector for precision and uniformity of outcomes and measures but is also visible in other fields (Lingard, 2011; Macdonald, 2013). Although the attention directed towards fitness levels and motor abilities is consistent with the accountability focus of the 'What Works' agenda, it may overly restrict the goals of physical education to its essential constituents. Emphasis on quantifiable goals, for instance, outcomes and qualifying levels, often leads to insensitivity to the relational, social, and emotional aspects of learning. Many would argue, we think, that the goal of PE is not only to promote ideal engagement in physical activities but also to engender a sustained interest in active lifestyles (Bailey, 2020). In addition, there is a real possibility that over-dependence on standardised instruments may discourage and thus exclude students who are outside the conventional level of sporting ability and create an environment which is unfriendly to lesser sporting people (Penney, 2013). The over-emphasis on objective outcomes, particularly in physical education, may unintentionally lead to valuing quick wins at the expense of individual and social growth in the long run and, in so doing, ignore the immense potential of physical education in shaping well-balanced, self-assured, and socially competent citizens.

Formalised assessment tools have become a fundamental feature of PE programmes because they allow for orderly approaches to measuring physical abilities and improvements in students over time. Most motor skill tests and standardised fitness tests evaluate skills such as endurance, power, and mobility and emphasise different dimensions of physical development in children, regardless of school. Such incorporation of standardised tools enhances the alignment of PE with an evidence-based approach that emphasises the use of strategies to improve the population's physical abilities and health (Bailey, 2006). Proponents of this approach maintain that establishing timetables for PE based on scientific procedures orientates the subject to the same level of precision and accountability as other curricula. This perspective reinforces its role as a worthwhile subject, leading to advantageous outcomes for learners (Alexander, 2008; Kirk, 2010).

The focus on physical health and competence shares commonality with increasing scientific literacy within the educational framework, as expressed through PE. In the same vein, with the advent of education systems, PE has increased and emphasised achievements such as improving cardiovascular fitness, motor coordination, and strength, all of which indicate the programme's success. These parameters guarantee uniformity and consistency, thus making it possible for all education providers to ensure that learners receive the same level of education in physically competent activities. Life skill proficiency and most national fitness tests provide cut-off points that teachers can use to measure students and guide their teaching and the education they intend to have (Williams & Lacy, 2018).

Physical educators who aim for ‘What Works’ point out that this approach increases the domain's authority by showing distinct and scientifically supported outcomes. As childhood inactivity and obesity are pressing problems worldwide, advocates claim that PE properly addresses student health and development in a measurable and recordable manner (Bailey, 2006; Blair & Cheng, 2018). Evidence-based PE methods explicitly aim to enhance physical skills, health, and other educational policies that emphasise effectiveness and targeted outcomes (Penney, 2013; Whitehead, 2010).

The performance evaluation approach has also attracted the attention of policy-makers and administrators, who seek straightforward and dependable methods for assessing the impact of education. As PE seeks uniformity of purpose and approaches, it fits into the data-oriented education approach, which emphasises the need for evidence of improvement (Lingard, 2011). This means that PE stands to benefit from expressing its inputs through quantitative outputs, which are part of the accountability culture dictated by educational policy. For instance, improvements in physical abilities and fitness gained through effective PE instruction count towards schools’ obligations to demonstrate programme effectiveness and the efficient utilisation of resources (Macdonald, 2013; Whitehead, 2010).

The movement in the centre of gravitation towards metrics may differ from the idea of a holistic, personal-centred, and thus custom-made development that is rarely associated with PE. Nonetheless, the focus on empirical, observable, and study-results-based outcomes has anchored PE in contemporary education. In addition, the ‘What Works’ approach provides impetus towards a consistent structure-led evidence orientation that enhances students' academic performance by addressing their physical and mental health (Blair & Cheng, 2018).

Connecting 'What Works' in Physical Education Research to Theory and Practice

In education, the question of "What Works" ultimately comes down to using evidence-based practices linked to existing theories such as constructivism, behaviourism, and social learning theory. Such instructional theories help understand how students learn and retain knowledge more effectively. One example is constructivism, which encourages learning through active engagement and activation of prior knowledge and has been successful for years in creating meaningful and long-lasting experiences (O’Connor, 2022). Behaviourism provides strategies for managing the class and, ideally, rewards positive actions by giving order to the learning environment (Hinduja, 2021). Social Learning Theory demonstrates how observational learning stems from peer interaction, reminding us that students constantly imitate the behaviours and attitudes they see in their peers and teachers (Rumjaun & Narod, 2020).

While these main theories underpin the discipline of PE, they have particular adaptations due to the nature of physical experience and learning. Students build motor skills and acquire cooperative behaviours through observation and mimicry, so Social Learning Theory is particularly relevant for PE (Rumjaun & Narod, 2020). The hands-on experiential learning principles associated with constructivism manifest in task-oriented activities with a strong element of exploration movement and skill practice, predominantly in a physical environment based on concepts formed in previous learning experiences (Dalkiran et al., 2020). Furthermore, Self-Determination Theory, which focuses on autonomy and competence, relates specifically to PE as students are more likely to engage in physical activity and adhere when they feel

a sense of independence and accomplishment (White et al.,2021). Integrating these theoretical learnings into practical content, PE can facilitate a more engaging, inclusive, and practically functional classroom that respects the recruitment of specialised learning processes within this domain as a stage for both skill development and broader life preparation.

Criticisms of 'What Works' in Physical Education

Experiencing or engaging in the 'What Works' agenda in PE has also drawn challenges. Although the emphasis on fitness and motor competence integrates PE in the larger accountability framework, it is believed to be a threat to the conventional view of PE, which is more than just encouraging these skills, as it also includes social, emotional, and motivational aspects. It has also been argued that in pursuing these standardised targets and technical skills, PE may just become a transfer of motor skills, which may not be very helpful in personal and social development. The inculcation of this target results in a reduction in the PE view to be target-orientated, meaning fitness targets rather than the development of appreciation towards PA (Houser & Kriellaars, 2023).

It has become increasingly claimed that 'physical literacy' can and should serve as an adequate measure for PE, although this is highly debated (Bailey, 2020). Whitehead's definition of physical literacy incorporates more than just the physical aspect; it includes the elements of motivation, confidence, and even cognitive engagement, intending to appreciate movement for a lifetime, which is positive. It is concerned with immersing PE students from learning motor techniques only to broader objectives centring on a person's development. However, Bailey (2020) views the application of physical literacy as rhetorically useful but virtually impossible to implement in practice. A more complex view of physical literacy would make it difficult to define it within an accountability paradigm, thus making it difficult to contextualise it within evidence-based education models.

Some physical educators are worried about the possible neglect of PA per se in favour of fitness and motor skill improvement in the 'What Works' approach. Like many other authors, Kretchmar (2000) emphasises that such a relationship should be established when students appreciate PA as something that can and should be done for the pleasure of it, not as a chore or an obligation. Kretchmar argues for an approach focused on "joy-oriented" development when play, search, and movement are the goals of education and are important in a broader educational context. Their criticism rests on the notion that by concentrating only on fitness and skill measures, teachers would estrange students who do not conform to the typical images of athletes, creating an environment centred on performance, which may hinder many students' participation (Oliver & Kirk, 2016). According to these critics, educators should strive to build students' physical capabilities, strengthen their emotional and social well-being, and foster feelings of cohesiveness and connectedness.

Another significant concern is the inflexibility of the 'What Works' method, which is central to PE. Most scholars agree that rigid, standardised learning cannot accommodate students' diverse abilities and interests (Macdonald, 2013). Since PE is based on socially and politically accepted norms, the focus on standardised assessment makes PE run the risk of not applying to a large group of students. Other educators have also opposed such models, as they may push students into aggression and competition on strategies to achieve certain goals and become more averse to physical engagement. Proponents of physical literacy, such as Whitehead (2010), have focused on such goals and have suggested broadening the vision to include motivation, confidence, and

personal involvement. In their view, true inclusiveness in PE requires engagement and variability, allowing teachers to respond to their practice in the specific context of their students. For Penney (2013), this position influences individualistic flexibility, which is the opposite of standardisation.

To sum up, critics of the 'What Works' move into PE suggest focusing on fitness and motor skills in PE might benefit systems and measurement of teachers, but it could limit the educational breadth of concepts within PE. They say that, with a focus on quantifiable objectives, PE will likely focus on the technical aspects of skill development and lose sight of the social, emotional, and motivational factors that enrich students' lives and inspire lifelong involvement in PA. Supporters of the broader view say that PE must be flexible, adaptable, and personalised, emphasising creativity and self-expression.

Rapprochement: Balancing Evidence-Based Practice with Holistic Goals

To minimise the impact of the criticisms recapped above, some scholars have taken a more moderate view to combine the best of the two halves of the 'What Works' model with the values usually associated with education in its broad sense. It aims to provide students with technical skills while addressing their social and emotional aspects. One such influential model is the Teaching Personal and Social Responsibility (TPSR) model developed by Hellison (2011). According to TPSR, physical skills are complemented with morals, such as caring and discipline, among other values. TPSR is also a curriculum framework aimed at helping teachers improve students' physical skills within educational values and ethical goals. This illustrates how PE teachers, when achieving ethical goals, develop moral values in their students by shaping their physical abilities. This PE model is undemocratic and highly stressful, and its main priority is educational value rather than violence.

Another successful strategy for reconciling the empirical and more holistic approaches in PE is mixed-methods research. This method presents a more comprehensive view of student participation in PE by focusing on quantitative statistics regarding students' physical capabilities and qualitative dimensions of their self and social aspects. This strategy upholds the empirical rigour behind defining the 'What Works' approach; it presents evidence of quantifiable results while accepting the intricacies of educational processes, which could be difficult to assess through standard metrics. For instance, student interviews, reflective journals, and observational data help provide PE's emotional and social aspects beyond mere technical competencies (Durden-Myers & Whitehead, 2018). In this context, the mixed method enables PE teachers and researchers to appreciate and uphold the multi-faceted nature of the learning process in PE so that skills and experiences are promoted.

The conception of PE as an active field of education lies in the professional autonomy and flexibility of the teachers. Many scholars argue that teachers should have the professional discretion to shift their focus to individual students' needs, interests, and backgrounds. Only by giving autonomy to teachers who exercise professional judgement can schools have a directive PE curriculum that processes evidential targets and advocates for a holistic and individualised integration of learning. This type of autonomy helps teachers incorporate evidence-based practices without loss of flexibility so that a supportive climate is created in which students are appreciated as unique individuals. Therefore, such an approach motivates students to have a healthy and active lifestyle throughout their lives aside from feeling competent and having a sense

of self that fosters attachment (Penney, 2013).

Conclusion

The 'What Works' agenda has brought significant advances in evidence-based practice, ensuring greater accountability and consistency in educational outcomes. However, its emphasis on measurable results risks narrowing the scope of education and overlooking the essential social, emotional, and developmental aspects. To achieve a meaningful balance, educational systems should integrate evidence-based practices with holistic approaches that prioritise creativity, autonomy, and the development of well-rounded individuals. This balanced perspective would improve measurable outcomes and foster lifelong learning, critical thinking, and personal growth, aligning education with its broader societal and ethical goals.

Funding

The authors were not in receipt of any funding for this study.

Disclosure statement

The authors report no potential competing interests.

Notes on Contributors

Richard Peter Bailey is a Professor of Education and Deputy Dean at UCSI University, Malaysia. He works at the interface between science and philosophy. Stanford University's list of top scientists in the world includes him for Clinical Medicine, Sport Sciences, and Education. Please direct correspondence to richardbailey@ucsiuniversity.edu.my

Nadia Samsudin is an accomplished academic and researcher at UCSI University. She specialises in applied statistics in public health and currently works as a post-doctoral research fellow in the Faculty of Social Sciences and Liberal Arts (FOSSLA). Her expertise and contributions play a crucial role in advancing knowledge and improving public health outcomes.
Please direct correspondence to nadia.samsudin@ucsiuniversity.edu.my

Hashem Salarzadeh Jenatabadi is a Senior Lecturer at the Department of Science and Technology Studies, Faculty of Science, University of Malaya, Malaysia. He holds Postdoc, PhD degrees in Applied Statistics, from University of Malaya, and master, bachelor's degrees in mathematical Statistics areas from Ferdowsi University of Mashhad, Iran. His expertise involves application of Statistical and Mathematical modeling in Engineering, Management, Public health, Chemistry, and Economics. Please direct correspondence to jenatabadi@um.edu.my

Rogayah A Razak is Dean of the Faculty of Social Science & Liberal Arts at UCSI University, Malaysia, and an expert in psycholinguistics and clinical linguistics. With over 35 years of experience in academia, she specializes in language development for both typically and atypically developing children, adults with aphasia, and the creation of language assessment tools for speech therapy. Please direct correspondence to rogayah@ucsiuniversity.edu.my

Francis Ries is an Associate Professor at the Faculty of Educational Sciences (University of Seville – Spain), affiliated to the Department of Physical Education and Sports. His current research interests focus on the impact of physical activity on mental health among university students and psychosocial aspects of physical activity. Moreover, he is interested in meaningful physical activity and physical education. Since 2018, he is the main researcher of the SPAMP (Sport and Physical Activity from a Multidisciplinary Perspective). Please direct correspondence to fries@us.es

References

Alexander, R. (2008). *Essays on pedagogy*. Routledge.

Bailey, R. (2006). Physical education and sport in schools: A review of benefits and outcomes. *Journal of School Health*, 76(8), 397-401. https://doi.org/10.1111/j.1746-1561.2006.00132.x

Bailey, R. (2020). Defining physical literacy: Making sense of a promiscuous concept. *Sport in Society*. https://doi.org/10.1080/17430437.2020.1777104

Bailey, R., Armour, K., Kirk, D., Jess, M., Pickup, I., Sandford, R., & BERA Physical Education and Sport Pedagogy Special Interest Group. (2009). The educational benefits claimed for physical education and school sport: An academic review. *Research Papers in Education*, 24(1), 1-27. https://doi.org/10.1080/02671520701809817

Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy, 18*(2), 215-228.

Baxter, J. & Clarke, J. (2013). Farewell to the tick box inspector? Ofsted and the changing regime of school inspection in England. *Oxford Review of Education*, 39(5), 702-718.

Biesta, G. (2007). Why "What Works" won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory, 57*(1), 1-22.

Biesta, G. (2010). *Good education in an age of measurement: Ethics, politics, democracy*. Routledge.

Blair, S. N., & Cheng, Y. (2018). Physical activity interventions in youth: Meeting the needs of the future. *Kinesiology Review, 7*(2), 101–106.

Carnine, D. (2000). *Why education experts resist effective practices (and what it would take to make education more like medicine).* Thomas B. Fordham Foundation.

Coalition for Evidence-Based Policy. (2002). *Bringing evidence-driven progress to education: A recommended strategy for the U.S. Department of Education*. Coalition for Evidence-Based Policy.

Coe, R., & Kime, S. (2019). *A (new) manifesto for evidence-based education: Twenty years on. Sunderland*. Evidence Based Education.

Dalkiran, O., Eryigit, F., & Sivri, S. (2020). Comparison of the Effects of Constructivist Learning on Cognitive, Affective and Psychomotor Fields Applied in Physical Education Courses. *African Educational Research Journal*, *8*, 327-334. https://doi.org/10.30918/AERJ.8S2.20.062

Davies, H. T. O., Nutley, S. M., & Smith, P. C. (Eds.). (1999). *What works?: Evidence-based policy and practice in public services*. Policy Press.

Dekker, I., & Meeter, M. (2022). Evidence-based education: Objections and future directions. Frontiers in Education, 7, 941410. https://doi.org/10.3389/feduc.2022.941410

Durden-Myers, E. J., & Whitehead, M. E. (2018). From philosophy to practice: Physical literacy in the 21st Century. *Journal of Teaching in Physical Education, 37*(3), 237-245. https://doi.org/10.1123/jtpe.2018-0136

Gorard, S. (2020). Getting evidence into education: Evaluating the routes to policy and practice. Routledge. https://doi.org/10.4324/9780429279088

Hargreaves, D. H. (1996). *Teaching as a research-based profession: Possibilities and prospects*. Teacher Training Agency.

Hellison, D. (2011). *Teaching personal and social responsibility through physical activity* (3rd ed.). Human Kinetics.

Hinduja, P. (2021). From behaviorism to constructivism in the teaching-learning process. *Journal of Education & Social Sciences*, *9*(2), 111-122. https://doi.org/10.20547/jess0922109204

HM Government. (2023). *What Works Network Strategy*. Evaluation Task Force. Retrieved from: https://www.gov.uk/government/organisations/evaluation-task-force

Houser, N., & Kriellaars, D. (2023). "Where was this when I was in Physical Education?" Physical literacy enriched pedagogy in a quality physical education context. *Frontiers in Sports and Active Living*, 5, 1185680.

Kirk, D. (2010). *Physical education futures*. Routledge.

Kretchmar, R. S. (2000). Joy and delight in movement. In R. L. Decker (Ed.), *Philosophy and education*. Stipes.

Lingard, B. (2011). Policy as numbers: Ac/counting for educational research. *Australian Educational Researcher, 38*(4), 355-382.

Macdonald, D. (2013). The physical education profession and its professional responsibility... or... why '12 weeks paid holiday' arguments don't wash. *Physical Education and Sport Pedagogy, 18*(5), 1-12.

Ndaruhutse, S., Jones, C., & Riggall, A. (2019). *Why systems thinking is important for the education sector*. Education Development Trust.

O'Connor, K. (2022). Constructivism, curriculum and the knowledge question: tensions and challenges for higher education. *Studies in Higher Education*, *47*(2), 412-422. https://doi.org/10.1080/03075079.2020.1750585

Oakley, A. (2000). *Experiments in knowing: Gender and method in the social sciences*. Polity Press.

Oakley, A., Gough, D., Oliver, S., & Thomas, J. (2005). The politics of evidence and methodology: lessons from the EPPI-Centre. *Evidence & Policy*, 1(1), 5-31.

OECD. (2004). *Learning for tomorrow's world: First results from PISA 2003*. Organisation for Economic Co-operation and Development.

Oliver, K. L., & Kirk, D. (2016). *Girls, gender and physical education: An activist approach*. Routledge.

Penney, D. (2013). Points of tension and possibility in physical education. *European Physical Education Review, 19*(1), 1-8. https://doi.org/10.1177/1356336X12465510

Pring, R. (2000). *Philosophy of Educational Research*. Continuum.

Pring, R. (2004). *Philosophy of education: Aims, theory, common sense, and research*. Continuum.

Rumjaun, A., & Narod, F. (2020). Social learning theory—Albert Bandura. *Science education in theory and practice: An introductory guide to learning theory*, 85-99. https://doi.org/10.1007/978-3-030-43620-9_7

Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher, 31*(7), 15-21.

Slavin, R. E. (2008). Perspectives on evidence-based research in education—What works? Issues in synthesizing educational program evaluations. *Educational Researcher, 37*(1), 5-14.

Thiedig, C. (2023). Evidence use in higher education decision-making and policy: A scoping review of empirical studies from 2010 to 2022. London Review of Education, *22*(1), https://doi.org/10.14324/LRE.22.36

White, R. L., Bennie, A., Vasconcellos, D., Cinelli, R., Hilland, T., Owen, K. B., & Lonsdale, C. (2021). Self-determination theory in physical education: A systematic review of qualitative studies. *Teaching and Teacher Education*, *99*, 103247. https://doi.org/10.1016/j.tate.2020.103247

Whitehead, M. E. (2010). *Physical literacy: throughout the lifecourse*. Routledge.

Williams, S. M., & Lacy, A. C. (2018). *Measurement and evaluation in physical education and exercise science*. Routledge.