Research Quarterly for Exercise and Sport

ISSN: 0270-1367 (Print) 2168-3824 (Online) Journal homepage: www.tandfonline.com/journals/urqe20

Is a Cup Enough? Levels of Energy Expenditure During Speed Stacking

Samuel C. Hymes, Ryan S. Sacko, Grace Duncan & Tim Bott

To cite this article: Samuel C. Hymes, Ryan S. Sacko, Grace Duncan & Tim Bott (2026) Is a Cup Enough? Levels of Energy Expenditure During Speed Stacking, Research Quarterly for Exercise and Sport, 97:1, 86-94, DOI: 10.1080/02701367.2025.2549064

To link to this article: https://doi.org/10.1080/02701367.2025.2549064

Published online: 29 Aug 2025.

RESEARCH QUARTERLY FOR EXERCISE AND SPORT
2026, VOL. 97, NO. 1, 86–94
https://doi.org/10.1080/02701367.2025.2549064

Routledge
Taylor & Francis Group

Is a Cup Enough? Levels of Energy Expenditure During Speed Stacking

Samuel C. Hymes, Ryan S. Sacko, Grace Duncan, and Tim Bott

The Citadel

ABSTRACT

This study evaluated whether Speed Stacking elicits sufficient physiological intensity to meet moderate-to-vigorous physical activity (MVPA) thresholds in youth, using portable indirect calorimetry and accelerometry. Fifty-four participants (*M*age = 14.01 ± 5.48 years; 46.3% female) completed four randomized activity conditions: Rest, Speed Stacking, Running, and Speed Stack + Run, each lasting five minutes. Participants wore a COSMED K5 metabolic system, hip and wrist ActiGraph GT3X+ accelerometers, and a heart rate monitor. Energy expenditure (METs), accelerometer counts, perceived exertion, and fun were analyzed using repeated measures ANOVAs and Cohen's Kappa for method agreement. In both prepubescent and pubescent groups, Speed Stacking alone did not meet the specific MET thresholds for MVPA. However, the hybrid Stack + Run condition significantly increased METs (up to 5.1 ± 1.2 in prepubescents). Cohen's Kappa revealed poor agreement between MET-based and accelerometer-based intensity categorical classifications across all conditions. Participants, especially younger children, reported Speed Stacking as fun but minimally effortful. Speed Stacking alone fails to meet key pedagogical criteria for physical education, particularly those related to maximizing activity time and MVPA engagement. Given the limited time available in PE settings, alternative activities with higher energy expenditure may offer a more efficient use of instructional time. However, when integrated with movement-based tasks, Speed Stacking can be adapted to better support physical and educational outcomes.

ARTICLE HISTORY
Received 12 May 2025
Accepted 13 August 2025

KEYWORDS
Energy expenditure; physical education curriculum; sport stacking; youth physical activity

In an era where physical inactivity has become a global public health concern, schools are increasingly viewed as a critical and accessible environment for promoting physical activity (PA) during childhood (Pate et al., 2006). For millions of children and adolescents, the school day represents the primary, and sometimes only, structured opportunity to engage in moderate-to-vigorous physical activity (MVPA (Brazendale et al., 2017). With limited time allocated to physical education (PE) in many school systems, every minute counts, and the selection of activities must reflect high-value opportunities, those that combine high student engagement, sufficient PA intensity, and positive affective experiences (i.e., fun and enjoyment), to maximize health benefits and learning outcomes. According to current PA guidelines established by the World Health Organization (WHO), children and adolescents aged five to 17 should accumulate at least 60 minutes of MVPA daily (Chaput et al., 2020). However, adherence remains alarmingly low, with only 16.1% of U.S. adolescents meeting these benchmarks (Nagata et al., 2022). This shortfall underscores the need for intentional, evidence-based approaches within PE curricula to ensure that the time allocated to PA translates into meaningful health outcomes. In addition to meeting PA guidelines, effective PE activities should align with established pedagogical frameworks that define quality learning experiences (Rink, 2020). Rink (2020) outlines four essential criteria for a good learning experience in PE: the potential to improve motor performance, the provision of maximal activity time, developmental appropriateness, and the integration of cognitive, affective, and psychomotor domains. Evaluating activities through this lens provides a more comprehensive understanding of their contribution to physical and educational outcomes. This framework offers a critical context for assessing Speed Stacking's value within PE curricula, beyond energy expenditure alone.

Physical Education classes, serve as critical venues for promoting daily MVPA. Traditionally, activities such as jogging, soccer, and tennis have been promoted as ideal ways for children to achieve their daily MVPA (Ainsworth et al., 2011; Butte et al., 2017). Yet, achieving the recommended MVPA intensity during PE can be challenging due to structural and curricular constraints (Saint-Maurice et al., 2018). National guidelines advocate that at least 50% of PE class time should be spent in MVPA, but not all activities integrated into PE curricula support this goal (America et al., 2014). One such activity is Speed Stacking (also known as sport stacking or cup stacking), a hand-eye coordination task that gained traction in PE settings over the past 20 years.

Organizations like SHAPE America have endorsed Speed Stacking for its motor skill benefits, citing its potential to enhance bilateral coordination and reaction time (Hart et al., 2006; Udermann et al., 2004). However, evidence supporting Speed Stacking related to its ability to promote minimal thresholds of MVPA remains scant, especially in children. Energy expenditure during PA can be equated using a unit of metabolic equivalent of task (MET), where 1.0 METs represents the caloric equivalent of sitting quietly and physical activities, classified as moderate intensity fall between 3.0 and 6.0 METs in adults and between 4.0 and 7.0 METs in children

CONTACT Ryan S. Sacko rsacko@citadel.edu Health and Human Performance, The Citadel, 171 Moultrie Street, Charleston, SC 29409.

(U.S. Department of Health and Human Services [HHS], 2018). A foundational study by Murray et al. (2009) using stationary indirect calorimetry to estimate METs during Speed Stacking reported an average energy expenditure of 3.1 METS, just meeting the lower bound of moderate-intensity threshold for adults (Murray et al., 2009). However, for children, the moderate-intensity threshold of ≥ 4.0 METs was not met (Ainsworth et al., 2011; Butte et al., 2018). However, compared to standard PE activities. many of which regularly elicit MET values in the range of 6.0–8.0 for children, Speed Stacking's lower energy demands raise questions about its contribution to MVPA goals (see Table 1).

Established technologies, such as portable indirect calorimetry systems (e.g., COSMED K5) and accelerometry, provide the opportunity to reassess Speed Stacking under more dynamic, ecologically valid conditions (Crouter et al., 2019). These tools offer improved data capture in free-living or classroom-like environments. Indirect calorimetry, particularly when conducted with portable systems, is widely regarded as the gold standard for assessing energy expenditure in real-world settings. Unlike stationary metabolic carts, portable calorimeters allow for natural movement and can provide breath-by-breath estimates of oxygen consumption and METs during activities. This level of ecological validity is particularly critical when evaluating activities such as Speed Stacking, where movement is dynamic and varied but may not consistently trigger large-muscle group activation (Sacko, Nesbitt, et al., 2019). While indirect calorimetry offers a criterion measure of energy expenditure, accelerometry remains a widely adopted method in school-based PA surveillance due to its practicality and capacity for continuous, non-invasive monitoring (Crouter et al., 2019; Sacko, Brazendale, et al., 2019; Sacko, Nesbitt, et al., 2019). In this study, accelerometers were not used for direct validation purposes; rather, they were used to explore their translational relevance. Specifically, whether movement data derived from accelerometers could meaningfully approximate the intensity of activity as measured by calorimetry. This dual-methods approach was motivated by the recognition that cut-points and regression equations used in accelerometry are often calibrated on locomotor activities and may not generalize to non-traditional, skill-based tasks such as Speed Stacking. As demonstrated in prior work involving object projection skills, accelerometers may significantly underestimate the energy expenditure of discrete, intermittent, and upper-limb focused activities, even when such activities elicit moderate-to-vigorous metabolic intensities (Sacko, Brazendale, et al., 2019; Sacko, McIver, et al., 2019). Accelerometers are commonly used to estimate time spent in MVPA based on predefined intensity cut points. However, these cut points are often derived from locomotor-dominant activities like walking or running and may not generalize well to non-traditional or skill-based tasks such as object manipulation or Speed Stacking, where energy expenditure may be under- or overestimated relative to the actual metabolic cost (Duncan et al., 2020; Trost, 2001; Trost et al., 2005). This methodological limitation can obscure the true intensity classification of such activities (Sacko, Brazendale, et al., 2019). Given these limitations, our inclusion of both methods was partially motivated by a desire to examine the congruence or divergence between objective calorimetric data and more practical, school-based tools. While not designed as a formal validation study, this study contributes exploratory insight into how accelerometer-derived estimates may align with indirect calorimetry.

Speed Stacking is not just a stationary activity, as measured by Udermann et al. (2004), it can be more dynamic when coupled with continuous activities associated with large muscle recruitment and higher intensity levels (Udermann et al., 2004). There has been a growing interest in hybridizing Speed Stacking with movement-based interventions to increase the intensity of the PA and enhance MVPA time. For example, integrating locomotor components (e.g., running between stacks) or pairing with cardiovascular tasks can enhance energy expenditure and student engagement. These hybrid game scenarios align with the broader pedagogical shift toward cognitively and physically integrated PE instruction, as outlined in recent frameworks for enriched PE environments (Granero-Gallegos, 2020; Hulteen et al., 2023; Li et al., 2011). Ultimately, reproducibility and rigor in assessing PA intensity are essential

Table 1. Estimated MET values for selected physical education activities by age group.

PE Activity	6 yr − 9 yr	10 yr − 12 yr	13 yr − 15 yr	16 yr − 18 yr	≥19 yr
Soccer Game	7.7	8.1	8.4	8.7	9.5
Bouncing, kicking, dribbling ball, reaction ball (moderate intensity)	6.0	6.2	6.3	6.5	–
Bouncing, kicking, dribbling ball, reaction ball (vigorous intensity)	6.1	6.3	6.4	6.6	–
Freeze/zone tag Moderate intensity	6.3	6.5	6.6	6.7	–
Freeze/zone tag Vigorous intensity	6.4	6.6	6.7	6.9	–
Jump rope	6.9	7.1	7.2	7.4	–
Rope jumping, fast pace, 120–160 skips/min	–	–	–	–	12.3
Rope jumping, moderate pace, 100–12 skips/min	–	–	–	–	11.8
Sharks and minnows	5.8	6.0	6.1	6.2	–
Jog—Fast	7.2	7.9	8.5	8.8	–
Jog—Slow	5.5	5.9	6.3	6.7	–
Run—6.0 mph	8.2	9.1	9.8	10.5	9.3
Run—8.0 mph	10.6	11.5	12.4	13.2	12.0
Basketball Game	6.7	7.0	7.2	7.5	8.0
Shooting and retrieving a basketball	5.9	6.2	6.4	6.6	–
Shooting baskets	–	–	–	–	5.0
Drills, practice	–	–	–	–	9.3

Note: MET = metabolic equivalent of task. Values represent estimated energy expenditure for common physical education activities across age groups. All MET values were derived from the *Youth Compendium of Physical Activities* and the *2011 Compendium of Physical Activities* (Ainsworth et al., 2011; Butte et al., 2018).

components and foundations of scientific research. While preliminary studies have supported the classification of Speed Stacking as a moderate-intensity activity (Murray et al., 2009), the inconsistency in measurement tools, participant demographics, and protocol design calls for further validation. Therefore, the purpose of this study is to replicate and extend the existing research by evaluating the energy expenditure of Speed Stacking, using portable indirect calorimetry as a criterion method, while also incorporating accelerometry to examine its potential utility as a practical, school-based alternative for estimating physical activity intensity. This inclusion enabled assessment of the degree of classification alignment between accelerometer outputs and calorimetric data, an issue of methodological relevance for field-based monitoring of non-locomotor, skill-based activities within physical education contexts. Findings from this study aim to inform PE practitioners of the utility of Speed Stacking within the broader context of MVPA-promoting curricular design.

Methods

Participants

Fifty-four participants (*M*age = 14.01. ± 5.48 years; 46.3% female) from the southeastern U.S. were voluntarily enrolled for the purposes of this study. Participants were generally healthy, with no known cardiovascular, metabolic, or orthopedic conditions. The study was approved by The Citadel's Institutional Review Board (IRB Committee, protocol 2223–29). Informed consent was obtained from all adult participants, and children provided informed assent, with parental consent also secured in accordance with IRB guidelines. Participants were excluded if they (1) were unable to complete a 3–6–3 stacking sequence (see Figure 1) in under 8.0 seconds (Murray et al., 2009), (2) consumed caffeine or exercised prior to testing, (3) had a positive response on the PAR-Q (2023) health screening, or (4) voluntarily withdrew at any time.

Figure 1. Participant performing a 3–6–3 Speed Stacking sequence using regulation cups.

Experimental design

Participants completed all four testing conditions: Baseline Rest, Speed Stacking, Running, and Speed Stack + Run, in a single session, each lasting five minutes. The Running and Speed Stack + Run conditions were included to reflect common PE locomotor activities and to evaluate whether the integration of stacking with ambulatory movement increases physiological intensity. Conditions were administered in a randomized, counterbalanced order to control order effects, except for the Baseline Rest condition, which always occurred first. All testing occurred in a gymnasium (hardwood basketball floor) to simulate a PE environment, and all environmental conditions (temperature ~22°C, humidity controlled) were held constant across sessions.

The Speed Stacking and Speed Stack + Run conditions followed guidelines from the World Sport Stacking Association (WSSA), including the use of regulation stacking cups and standardized table height (28–32 inches). To simulate the timing and rhythm of a typical PE relay activity, the Running and Speed Stack + Run conditions used a 40-second interval. The interval was calculated based on the approximate time required to (1) run to the stacking table positioned seven feet away, (2) complete a 3–6–3 stacking sequence (average time: eight seconds), (3) return to the starting line, and wait while three hypothetical teammates completed the same sequence (see Figure 2). Prior to testing, participants completed a stacking skills assessment to verify eligibility by ensuring that they could perform the 3–6–3 sequence in under 8.0 seconds. Participant height and weight were measured using a mechanical physical beam scale.

Each participant wore a COSMED K5 portable metabolic system (via facemask), two ActiGraph GT3X+ accelerometers (on the right hip and non-dominant wrist), and a chest-worn Garmin HRM-Dual heart rate monitor. Data from the final two minutes of each condition were analyzed to capture steady-state energy expenditure. A standardized audio recording was used throughout the testing protocol to guide participants and maintain consistency across trials. The recording included: (1) pre-condition instructions, (2) verbal start cues, (3) one-minute remaining cues, and (4) completion prompts. After each activity, participants were seated until their heart rate returned to resting baseline values.

Baseline Rest condition

Participants stood still, breathed normally, and refrained from speaking or moving for five minutes. This condition served to

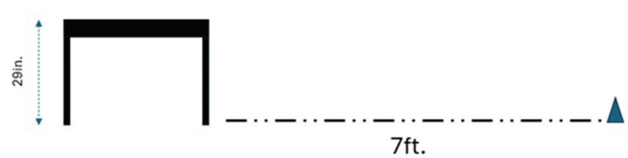

Figure 2. Diagram of testing setup showing table position and seven-foot run distance.

establish baseline metabolic and heart rate values. During the condition, participants followed audio cues to mark the start, time remaining, and completion.

Speed Stacking condition

Participants stood at a regulation-height folding table (29 inches) and completed as many 3–6–3 stacking cycles as possible within five minutes. Each cycle required sequentially upstacking and downstacking of cups in the 3–6–3 configuration. Trained research team members ensured proper sequence and technique throughout. Upon completion of the condition, participants sat quietly in a chair until their heart rate returned to baseline.

Running Condition

Participants ran from a designated start point to the stacking table (seven feet), touched the table, and returned, repeating this pattern every 40 seconds, simulating a relay race. Participants rested passively between repetitions. Audio cues provided start signals and time notification. Heart rate recovery was monitored post-condition.

Speed Stack + Run condition. This hybrid condition combined the previous two tasks. Participants ran from the start line to the stacking table (seven feet), completed one full 3–6–3 stacking cycle, then returned to the start line. This sequence was repeated every 40-seconds for five minutes.

Instrumentation and data collection

Accelerometry data were collected using ActiGraph GT3X+ accelerometers (ActiGraph, Pensacola, FL), worn on the right anterior hip and non-dominant wrist. Devices were initialized with a sampling rate of 100 Hz, and data were downloaded and processed using ActiLife6 software (v6.13.6 Pensacola, FL) using a 1-second epoch and vector magnitude counts. Activity classification was based on validated hip (Evenson et al., 2008; Freedson et al., 2005; Sasaki et al., 2011) and non-dominant wrist (Chandler et al., 2016) cut points.

Oxygen consumption (VO_2) was measured via indirect calorimetry using the portable COSMED K5 to collect expired respiratory gases on a breath-by-breath basis (VO2 $kg^{-1}\cdot min^{-1}$) and averaged over the final two minutes of each condition. METs were calculated by dividing VO_2 by 3.5 mL/kg/min. The COSMED K5 unit was worn according to manufacturer guidelines, with all harness and facemask fittings adjusted for comfort and seal integrity. The system was calibrated prior to each testing session using certified gas mixtures and a volume calibration syringe, in accordance with standard operating procedures (Perez-Suarez et al., 2018).

Children's Rate of Perceived Exertion (RPE) Scale

A validated 10-point pictorial scale was used to measure subjective physical exertion. The scale ranged from 1 "not tired at all") to 10 ("so tired I can't go anymore"). Participants were instructed to self-rate their physical effort immediately before and after each activity condition. Verbal and visual anchors were provided to ensure comprehension, particularly for younger participants. The scale format was adapted from the Children's OMNI Scale of Perceived Exertion (Utter et al., 2002), which builds on the original Borg framework (Borg, 1998) to provide an age-appropriate method for assessing perceived exertion in pediatric populations.

Fun scale

A 10-point affective scale, modeled after the RPE scale, was used to assess perceived fun of each activity. Ratings ranged from 1 ("not fun at all") to 10 ("the most fun I've ever had"). Participants rated their level of fun prior to and immediately following each condition. The scale was selected for its ease of administration, real-time applicability, and age-appropriate design, with clear face validity for the PE setting.

Data analysis

Participants were stratified into prepubescent (≥11 years; $n = 22$) and pubescent (≥12 years; $n = 32$) groups. One-way repeated measures ANOVA were used to examine differences in METs across the four activity conditions (Rest, Stack, Run, and Run + Stack). When significant main effects were identified, pairwise comparisons were conducted using Bonferroni-adjusted post hoc tests. Repeated measures ANOVAs were also used to examine METs and accelerometer counts across the four activity conditions as well as changes in RPE and Fun ratings across activity conditions. Additionally, Pearson's correlations were conducted to explore relationships between METs, RPE, and Fun across conditions. Raw MET values and accelerometer counts were first transposed into categorical classifications of activity intensity (e.g., light, moderate, vigorous) using established cut-points in Microsoft Excel. Notably, these cut-points were derived from tasks that did not include Speed Stacking among the original calibration activities. This allowed for a direct comparison of intensity classifications across modalities (See Table 2 for cut-point thresholds). To evaluate the agreement between categorical classification of

Table 2. Vertical axis cut-points associated with moderate-to-vigorous physical activity.

		Range of accelerometer counts-per-minute				
Cut-point	Wear location	Sedentary	Light	Moderate	Vigorous	Very-Vigorous
[a]Chandler et al	Wrist (non-dominant)	0–1932	1933–6348	6349–17532	> 17554	N/A
Evenson et al	Hip	0–100	101–2295	2296–4011	> 4012	N/A
[b]Freedson et al	Hip	0–149	150–499	500–3999	4000–7599	> 7600
[c]Freedson et al	Hip	–	0–2689	2690–6166	6167–9642	> 9643

Note: All cut-points are presented in counts-per-minute (CPM) and reflect vertical axis measurements.
[a]Originally published as counts per 5 seconds. N/A = not applicable.
[b]Freedson et al. (2005) and is intended for prepubescent children.
[c]Freedson et al., (2011) and is intended for pubescent individuals.

activity intensity (e.g., light vs. moderate), Cohen's Kappa coefficients were calculated to assess consistency between MET-based and accelerometer-based classifications across conditions. Statistical significance was set at $\alpha < .05$. Statistical analyses were conducted using SPSS Version: 29.0.0.0 (241) (IBM Corp., Armonk, NY).

Results

Demographic and descriptive statistics for each age group are presented in Tables 3 and 4. Across the four activity conditions (Rest, Stack, Run, and Run + Stack), descriptive means and standard deviations are reported for METs, hip and wrist accelerometer counts, RPE, and Fun scores.

In the prepubescent group, MET values ranges from 2.5 ± 0.4 at rest to 5.1 ± 1.2 during the Run + Stack condition. In the pubescent group, METs ranged from 1.7 ± 0.6 at rest to 3.9 ± 1.2 during the Run + Stack condition. Hip and wrist accelerometer counts increased progressively with activity intensity in both groups, with peak values observed in the Run + Stack condition. RPE values remained low across all conditions for both groups, while Fun scores were higher in the prepubescent group compared to the pubescent group across all activity types.

A one-way repeated measures ANOVA revealed a significant main effect of activity condition on METs in both prepubescent ($F(3,18) = 58.31$, $p < .001$, $\eta^2 = .91$) and pubescent participants ($F(3,29) = 105.71$, $p < .001$, $\eta^2 = .92$), indicating that energy expenditure differed significantly across conditions. Bonferroni-adjusted post hoc tests revealed that MET values for the Run + Stack condition were significantly greater than all other activity conditions ($p < .01$), in both prepubescent and pubescent groups. Stack and Run also differed significantly from Rest ($p < .001$), and Stack differed significantly from Run in both groups ($p < .05$).

Significant main effects of activity condition were also found for hip accelerometer counts in both the prepubescent ($F(3,17) = 115.56$, $p < .001$, $\eta^2 = .95$) and pubescent groups ($F(3,23) = 111.23$, $p < .001$, $\eta^2 = .94$), as well as for wrist counts in prepubescent ($F(3,17) = 251.15$, $p < .001$, $\eta^2 = .98$) and pubescent participants ($F(3,23) = 334.54$, $p < .001$, $\eta^2 = .98$). Bonferroni-adjusted post hoc tests indicated that the Run + Stack and Run conditions elicited significantly greater hip accelerometer counts than Stack and Rest ($p < .001$). However, the difference between Run and Run + Stack was not statistically significant ($p > .05$). For wrist accelerometry, all activity conditions significantly differed from one another ($p < .01$), with Stack yielding the highest mean wrist counts.

For RPE, a significant main effect of condition was observed in the prepubescent group ($F(3,23) = 3.13$, $p = .045$, $\eta^2 = .29$), however, the effect was not statistically significant in the pubescent group ($F(3,17) = 2.78$, $p = .73$). A similar trend was observed for Fun scores. A significant main effect of condition was found in the prepubescent group ($F(3,23) = 4.40$, $p = .014$, $\eta^2 = .37$). No significant effect was observed in the pubescent group ($F(3,17) = 3.05$, $p = .057$), although the trend approached significance.

Cohen's Kappa coefficients were calculated to assess the level of agreement between MET-derived MVPA classification (≥4.0 METs in children, ≥ 3.0 METs in adults) and those based on established accelerometer cut points (i.e., Freedson, Evenson, and Chandler) across all activity conditions (Stack, Run, Run + Stack). Across all comparisons (see Table 5), no statistically significant agreement was observed (all p-values ≥ .05). In the Stack, Run, and Run + Stack activity conditions, none of the cut point schemes (i.e., Freedson, Evenson, or Chandler) produced significant agreement in either prepubescent or pubescent groups. Several comparisons resulted in constant classifications (i.e., all data points fell into a single category), particularly

Table 3. Descriptive statistics for physiological, perceptual, and affective measures across activity conditions in prepubescent participants.

		Rest	Stack	Run	Run + Stack
N = 22(F = 11)	METS	2.5 ± 0.4	3.6 ± 1.0	4.5 ± 1.2	5.1 ± 1.2
Age 8.9 ± 1.4	Hip (counts)	2.8 ± 4.2	78.9 ± 96.2	381.2 ± 104.3	466.2 ± 139.7
Height 139.6 ± 10.5	Wrist (counts)	220.8 ± 256.7	4605.1 ± 865.0	2371.2 ± 953.3	3861.6 ± 1335.2
Weight 34.0 ± 9.7	RPE	2.0 ± 1.3	2.8 ± 1.4	3.0 ± 2.1	2.8 ± 1.7
Mean Complete Stacks 16.0 ± 4.1	FUN	5.9 ± 2.7	7.7 ± 1.7	6.7 ± 2.9	7.0 ± 2.4

Note: Pubescent participants (≥12 years old). METs = metabolic equivalents; RPE = rate of perceived exertion (10-point pictorial scale); Fun = participant-rated enjoyment (10-point scale). Accelerometer counts: Hip values classified using Freedson et al. (2005) cut points; non-dominant wrist values classified using Chandler et al. (2016) cut points.

Table 4. Descriptive statistics for physiological, perceptual, and affective measures across activity conditions in pubescent participants.

		Rest	Stack	Run	Run + Stack
N = 32(F = 16)	METS	1.7 ± 0.6	2.9 ± 0.5	3.5 ± 1.4	3.9 ± 1.2
Age 17.8 ± 3.7	Hip (counts)	5.9 ± 10.2	46.9 ± 69.3	449.2 ± 135.7	525.4 ± 149.3
Height 164.3 ± 11.7	Wrist (counts)	229 ± 275.0	5189.3 ± 860.1	2101.4 ± 1100.1	3568.7 ± 960.9
Weight 65.5 ± 17.4	RPE	2.1 ± 1.4	2.4 ± 1.5	2.5 ± 1.5	2.8 ± 1.5
Complete Stacks 24.0 ± 5.2	FUN	4.9 ± 2.6	5.4 ± 2.5	4.8 ± 2.6	5.6 ± 2.6

Note: Pubescent participants (≥12 years old). METs = metabolic equivalents; RPE = rate of perceived exertion (10-point pictorial scale); Fun = participant-rated enjoyment (10-point scale). Accelerometer counts: Hip values classified using Evenson (2008) child cut points; Wrist values classified using Chandler et al. (2016) cut points

Table 5. Cohen's Kappa coefficients comparing MET-Based and accelerometer-based MVPA classification across activity conditions and age.

Condition	Cut Point Scheme	Age Group	*Kappa*	*p*-value	Interpretation
Stack	Freedson	Prepubescent	−.07	.18	No agreement
		Pubescent	.00	–	Constant
	Evenson	Prepubescent	−.04	.36	No agreement
		Pubescent	−.08	.20	No agreement
	Chandler	Prepubescent	.20	.09	Slight agreement (NS)
		Pubescent	0	–	Constant
Run	Freedson	Prepubescent	−.07	.18	Slight agreement (NS)
		Pubescent	.00	–	No agreement
	Evenson	Prepubescent	−.04	.20	No agreement
		Pubescent	0	–	Constant
	Chandler	Prepubescent	−.08	.19	No agreement
		Pubescent	−.03	.57	No agreement
Run + Stack	Freedson	Prepubescent	.24	.10	Slight agreement (NS)
		Pubescent	.00	–	Constant
	Evenson	Prepubescent	0	–	Constant
		Pubescent	0	–	Constant
	Chandler	Prepubescent	0	–	Constant
		Pubescent	−.04	.26	No agreement

Note: MET = metabolic equivalent of task; MVPA = moderate-to-vigorous physical activity; NS = not significant. Cohen's Kappa (*κ*) was used to evaluate agreement between MET-derived intensity classification (≥4.0 METs) and accelerometer-based MVPA classifications using four cut point schemes: Freedson et al. (2005) prepubescent, Freedson 2011 (Sasaki et al., 2011) pubescent, Evenson et al. (2008), and Chandler et al. (2016). Values represent Kappa coefficients with corresponding *p*-values. Interpretations follow Landis and Koch (1977): < 0 = no agreement, .00–.20 = slight, .21–.40 = fair, .41–.60 = moderate, .61–.80 = substantial, .81–1.00 = almost perfect. "Constant" indicates conditions in which all classifications fell into a single category, precluding statistical analysis.

among pubescent participants, which precluded statistical analysis. These findings suggest limited concordance between MET-derived and accelerometer-based MVPA classification methods under the tested conditions, regardless of age group or cut point scheme.

Discussion

The purpose of this study was to evaluate the energy expenditure and PA intensity of Speed Stacking using indirect calorimetry and accelerometry with particular attention to determining whether Speed Stacking, when performed independently, elicits sufficient physiological intensity to meet the youth-specific threshold for MVPA. Our findings demonstrate that Speed Stacking, when performed in isolation, does not consistently reach the child-specific 4.0 MET threshold for moderate intensity activity. In contrast, the hybrid Run + Stack condition elicited higher energy expenditures that approached or exceeded MVPA thresholds, particularly among pubescent participants.

These findings reaffirm prior research that questioned the physical intensity of Speed Stacking when conducted in a static manner (Murray et al., 2009). While earlier work reported MET values at or near the adult moderate threshold of 3.0 METs, our use of a portable metabolic system in an ecologically valid PE-like environment revealed values consistently below the recommended intensity for children, aligning with the youth-specific standards established by Butte et al. (2018). This highlights the need for age-appropriate thresholds when evaluating school-based activities.

While these physiological findings align with previous concerns about Speed Stacking's limited contribution to MVPA (Murray et al., 2009), integrating the four criteria (see Table 6) for effective PE learning experiences from Rink (2020) allows for a broader, educationally grounded interpretation of Speed Stacking's role in PE. When assessed against these criteria, motor performance potential, maximal activity time, developmental appropriateness, and cognitive-affective-psychomotor integration, Speed Stacking, as a stand-alone activity, demonstrates both strengths and several limitations.

Speed Stacking supports bilateral coordination, fine motor skills, and hand-eye coordination, competencies that are valuable for motor skill development. These attributes are particularly valuable for younger students or those with less developed gross motor skills, offering an accessible entry point for skill-building within diverse PE classrooms. However, its lack of gross motor engagement limits its effectiveness in promoting broader motor competence necessary for lifelong PA participation. Further, our data showed that Stack + Run and Run conditions produced comparable lower-body activity, as evidenced by similar hip accelerometer counts, underscoring the importance of locomotor components in elevating whole-body

Table 6. Alignment of Speed Stacking with the four criteria for good learning experiences in PE.

Criterion	Alignment with Speed Stacking (Stand-Alone)	Alignment with Run + Stack Hybrid
Motor Performance Improvement	Emphasizes bilateral coordination and fine motor skills; lacks gross motor challenge	Incorporates locomotor and gross motor engagement, promoting broader motor competence
Maximal Activity Time	Does not consistently meet child-specific MVPA threshold; low overall energy expenditure	Meets or exceeds youth MVPA threshold; promotes sustained physical engagement
Developmentally Appropriate	Best suited for prepubescent learners; perceived as less enjoyable and challenging by older youth	Suitable across age groups; intensity and variation support multilevel engagement
Integrates Cognitive, Affective, and Psychomotor	High enjoyment reported; limited psychomotor intensity and cognitive demands	Enhanced engagement across domains via movement integration and decision-making tasks

physical exertion. This reinforces the need to pair upper-limb focused tasks with dynamic movement to enhance motor competence and energy expenditure. Maximizing active engagement time is a central tenet of effective PE (Rink, 2020), particularly given national guidelines that advocate for at least 50% of PE class time spent in MVPA (National Physical Education Standards-SHAPE America Sets the Standards, 2014). Data indicates that Speed Stacking does not consistently elevate energy expenditure to MVPA levels, falling short of this criterion. In contrast, the Stack + Run hybrid condition significantly improved energy expenditure, suggesting that task design modifications that integrate locomotor components may help align Speed Stacking with this standard. When contextualized against other PE activities (listed in Table 1), this finding becomes more consequential. For example, even moderate forms of tag, ball-handling, or jumping rope routinely elicit MET values above 6.0 across all child and adolescent age groups, substantially exceeding the energy expenditure of Speed Stacking. Activities such as fast-paced jogging or soccer surpass 7.0–8.0 METs, underscoring how conventional PE tasks offer a more robust contribution to daily MVPA accumulation. Given these stark differences, allocating PE time to unmodified Speed Stacking should be done with clear intent and likely in complement, not replacement, of more metabolically demanding tasks.

Rink (2020) stresses the need for developmental appropriateness in learning tasks, considering the physical, cognitive, and emotional readiness of students (Rink, 2020). Our findings revealed that while prepubescent participants rated Speed Stacking as enjoyable and effortful, pubescent students reported lower enjoyment and perceived challenge. This age-dependent variability suggests that Speed Stacking may be best suited for younger learners or as a differentiated option within a multilevel instructional design, especially for students who may be less engaged by traditional sport-based or high-intensity activities. In this way, Speed Stacking can help support inclusive PE practices by accommodating varied interests and ability levels.

Effective PE experiences should foster learning across cognitive, affective, and psychomotor domains (Rink, 2020). Speed Stacking provides clear affective benefits as participants consistently rated the activity as enjoyable. However, in its basic form, Speed Stacking lacks the psychomotor intensity and decision-making demands that characterize higher-order learning experiences. This observation is consistent with wrist accelerometry data, which showed elevated upper-limb activity during Speed Stacking, suggesting its emphasis on fine motor engagement over full-body exertion. While this affirms its utility for targeting fine motor skill development, it also exposes the limited engagement of large-muscle groups critical for achieving moderate-to-vigorous intensity levels. The integration of movement-based or cognitively enriched modification, such as relays, problem-solving stacking challenges, or hybrid Stack + Run formats, could enhance this multidimensional learning potential. These designs call for enriched PE environments that simultaneously promote movement, engagement, and cognitive challenge (Hulteen et al., 2023). However, it is important to note that many traditional PE activities already embed higher-order cognitive and neuromotor demands through open-skill environments, exploratory movement, and tactical decision-making. Given these differences, Speed Stacking is best viewed as a supplementary or station-based activity rather than a substitute for more comprehensive motor learning experiences.

In terms of measurement, a notable discrepancy emerged between MET-derived intensity classifications and those derived from standard accelerometer cut points. Consistent with works from Sacko and colleagues, poor agreement was observed between these methods, particularly in the non-locomotor, skill-based task of Speed Stacking (Sacko, Brazendale, et al., 2019; Sacko, Nesbitt, et al., 2019). This lack of alignment calls attention to the limitations of accelerometry in accurately reflecting energy expenditure during manipulation-dominant tasks and reinforces the importance of employing content-sensitive tools in school-based PA assessment (Trost et al., 2005).

In addition to objective discrepancies in measurement tools, subjective perceptions of exertion and fun offered further insight into the multifaceted attributes of Speed Stacking as a PE activity. Participant perceptions of effort and enjoyment offered an important complementary perspective. RPE scores were low across all conditions, indicating that participants perceived Speed Stacking as minimally effortful. Interestingly, despite its low physiological load, Speed Stacking was rated as moderately fun, particularly among younger participants. These findings echo prior assertions that enjoyment and low perceived effort may bolster self-efficacy and PA adherence in PE settings, even if the activity itself does not substantially contribute to MVPA totals (Saint-Maurice et al., 2018).

Limitations

The sample was regionally specific to the Southeastern United States and predominantly composed of youth and young adults, which may limit the generalizability of findings across broader demographic populations. While sample size supported within-subject comparisons, larger, more diverse samples would improve external validity. The duration of each activity condition was limited to five minutes. Although appropriate for evaluating acute energy expenditure, this timeframe may not reflect how Speed Stacking contributes to cumulative MVPA over the course of a full PE class period or academic day. Furthermore, only the 3–6–3 stacking pattern and seven-foot run distance were evaluated, which may not represent the diversity of the Speed Stacking activities possible in classrooms. While indirect calorimetry via the COSMED K5 provides a gold standard measurement of energy expenditure, potential discomfort from wearing the device, particularly among younger children, could have influenced natural movement patterns. Additionally, accelerometer cut points applied in this study were developed predominantly for locomotor activities and may have systematically misclassified intensity for upper-body or static tasks. Notably, while youth-specific MET thresholds were employed to define moderate intensity, this approach does not capture individual variation in aerobic capacity. Personalized intensity zones derived from ventilatory thresholds would offer greater precision but were beyond the scope and feasibility of this study's design. Finally, although RPE and fun scores offer valuable insight into affective responses, younger children may have difficulty interpreting these consistently, despite efforts to provide visual and verbal anchors.

Conclusions

Speed Stacking, when implemented in isolation, does not reliably meet the moderate-intensity MVPA thresholds for youth, calling into question its efficacy as a stand-alone MVPA-promoting activity in physical education. However, hybrid models that integrate locomotor movement, such as Run + Stack substantially increase energy expenditure and better align with PE intensity standards set by WHO, Centers for Disease Control (CDC), and SHAPE America. These findings underscore the importance of using age-specific MET thresholds and highlight the limitations of accelerometer-based classification in non-locomotor tasks. They also demonstrate the value of triangulating physiological, perceptual, and affective data to assess PA holistically in educational contexts. While Speed Stacking alone may not suffice as a primary MVPA driver, it holds pedagogical value as a cognitively engaging, motor skill-enhancing, and enjoyable activity. When used strategically as part of warm-ups, stations, or hybrid modules, Speed Stacking can complement more vigorous activities and contribute meaningfully to inclusive, standards-based PE curricula.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

No funding was received for the conduct of this research or the preparation of this article.

ORCID

Ryan S. Sacko http://orcid.org/0000-0002-2104-6265

References

Ainsworth, B. E., Haskell, W. L., Herrmann, S. D., Meckes, N., Bassett, J. D. R., Tudor-Locke, C., Leon, A. S., Vezina, J., Whitt-Glover, M. C., & Leon, A. S. (2011). Compendium of physical activities: A second update of codes and MET values. *Medicine and Science in Sports and Exercise*, *43*(8), 1575–1581. https://doi.org/10.1249/MSS.0b013e31821ece12

Borg, G. (1998). *Borg's perceived exertion and pain scales*. Human Kinetics.

Brazendale, K., Beets, M. W., Weaver, R. G., Pate, R. R., Turner mcgrievy, G. M., Kaczynski, A. T., Chandler, J. L., Bohnert, A., & von Hippel, P. T. (2017). Understanding differences between summer vs. school obesogenic behaviors of children: The structured days hypothesis. *The International Journal of Behavioral Nutrition and Physical Activity*, *14*(1). https://doi.org/10.1186/s12966-017-0555-2

Butte, N. F., Watson, K. B., Ridley, K., Zakeri, I. F., McMurray, R. G., Pfeiffer, K. A., Crouter, S. E., Herrmann, S. D., Bassett, D. R., Long, A., Berhane, Z., Trost, S. G., Ainsworth, B. E., Berrigan, D., & Fulton, J. E. (2018). A youth compendium of physical activities: Activity codes and metabolic intensities. *Medicine and Science in Sports and Exercise*, *50* (2), 246–256. https://doi.org/10.1249/MSS.0000000000001430/

Butte, N. F., Watson, K. B., Ridley, K., Zakeri, I. F., McMurray, R. G., Pfeiffer, K. A., Long, A., Herrmann, S. D., Bassett, D. R., Long, A., Berhane, Z., Trost, S. G., Ainsworth, B. E., Berrigan, D., & Fulton, J. E. (2017). A youth compendium of physical activities: Activity codes and metabolic intensities. *Medicine and Science in Sports and Exercise*, *50* (2), 246. https://doi.org/10.1249/MSS.00000000000014300

Chandler, J., Brazendale, K., Beets, M., & Mealing, B. (2016). Classification of physical activity intensities using a wrist-worn accelerometer in 8-12-year-old children. *Pediatric Obesity*, *11*(2), 120–127. https://doi.org/10.1111/ijpo.12033

Chaput, J. P., Willumsen, J., Bull, F., Chou, R., Ekelund, U., Firth, J., Jago, R., Ortega, F. B., & Katzmarzyk, P. T. (2020). 2020 WHO guidelines on physical activity and sedentary behaviour for children and adolescents aged 5-17 years: Summary of the evidence. *The International Journal of Behavioral Nutrition and Physical Activity*, *17* (1). https://doi.org/10.1186/s12966-020-01037-z

Crouter, S. E., LaMunion, S. R., Hibbing, P. R., Kaplan, A. S., Bassett, D. R., & Loenneke, J. P. (2019). Accuracy of the Cosmed K5 portable calorimeter. *PLOS ONE*, *14*(12), e0226290. https://doi.org/10.1371/journal.pone.0226290

Duncan, M., Dobell, A., Noon, M., Clark, C., Roscoe, C., Faghy, M., Stodden, D. F., Sacko, R. S., & Eyre, E. (2020). Calibration and cross validation of accelerometry for estimating sedentary behavior and physical activity involving fundamental movement skills in children aged 8-12 years. *Sensors*, *20*(10), 2776. https://doi.org/10.3390/s20102776

Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., & McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *Journal of Sports Sciences*, *26*(14), 1557–1565. https://doi.org/10.1080/02640410802334196

Freedson, P., Pober, D., & Janz, K. F. (2005). Calibration of accelerometer output for children. *Medicine and Science in Sports and Exercise*, *37* (11), S523. https://doi.org/10.1249/01.mss.0000185658.28284.ba

Granero-Gallegos, A. (2020). New developments in physical education and sport. *International Journal of Environmental Research and Public Health*, *17*(24), 1–9. MDPI. https://doi.org/10.3390/ijerph17249171

Hart, M. A., Smith, L. A., & DeChant-Bruenning, A. (2006). Effect of participation in a cup stacking unit on hand-eye coordination of elementary children. *Physical Educator*, *3*(63), 869–876. https://doi.org/10.2466/pms.101.3.869-876

HHS. (2018). *2018 physical activity guidelines advisory committee scientific report*.

Hulteen, R. M., Terlizzi, B., Abrams, T. C., Sacko, R. S., De Meester, A., Pesce, C., & Stodden, D. F. (2023). Reinvest to assess: Advancing approaches to motor competence measurement across the lifespan. *Sports Medicine*, *53*(1), 33–50. https://doi.org/10.1007/s40279-022-01750-8

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159–174.

Li, Y., Ransdell, D., Irwin, L., Coleman, C., & Irwin, C. (2011). Sport stacking activities in school children's motor skill development. *Perceptual and Motor Skills*, *113*(2), 431–438. https://doi.org/10.2466/05.06.10.11.25.PMS.113.5.431-438

Murray, S. R., Udermann, B. E., Reineke, D. M., & Battista, R. A. (2009). Energy expenditure of Sport Stacking. *Physical Educator*, *66*(4), 108–185.

Nagata, J. M., Cortez, C. A., Dooley, E. E., Iyer, P., Ganson, K. T., & Pettee Gabriel, K. (2022). Moderate-to-vigorous intensity physical activity among adolescents in the USA during the COVID-19 pandemic. *Preventive Medicine Reports*, *25*. https://doi.org/10.1016/j.pmedr.2021.101685

Pate, R. R., Davis, M. G., Robinson, T. N., Stone, E., McKenzie, T., & Young, J. C. (2006). A leadership role for schools: A scientific statemeAmerican Heart Association Council on nutrition, physical activity, and metabolism (physical activity committee) in collaboration with the councils on cardiovascular disease in the young and cardiovascular nursing. *Circulation*, *114*(11), 1214–1224. https://doi.org/10.1161/CIRCULATIONAHA.106.177052

Perez-Suarez, I., Martin-Rincon, M., Gonzalez-Henriquez, J. J., Fezzardi, C., Perez-Regalado, S., Galvan Alvarez, V., Calbet, J. A., Morales-Alamo, D., & Calbet, J. A. L. (2018). Accuracy and precision of the COSMED K5 portable analyzer. *Frontiers in Physiology*, *9*, 1764. https://doi.org/10.3389/fphys.2018.01764

Rink, J. E. (Ed.). (2020). Designing learning experiences and tasks. In *Teaching physical education for learning* (8th ed., pp. 36–59). McGraw-Hill.

Sacko, R. S., Brazendale, K., Brian, A., McIver, K., Nesbitt, D., Pfeifer, C., & Stodden, D. F. (2019). Comparison of indirect calorimetry- and accelerometry-based energy expenditure during object project skill performance. *Measurement in Physical Education and Exercise Science*, *23*(2), 148–158. https://doi.org/10.1080/1091367X.2018.1554578

Sacko, R. S., McIver, K., Brazendale, K., Pfeifer, C., Brian, A., Nesbitt, D., & Stodden, D. F. (2019). Comparison of indirect calorimetry- and accelerometry-based energy expenditure during children‘s discrete skill performance. *Research Quarterly for Exercise and Sport*, *90*(4), 629–640. https://doi.org/10.1080/02701367.2019.1642440

Sacko, R. S., Nesbitt, D., McIver, K., Brian, A., Bardid, F., & Stodden, D. F. (2019). Children’s metabolic expenditure during object projection skill performance: New insight for activity intensity relativity. *Journal of Sports Sciences*, *37*(15), 1755–1761. https://doi.org/10.1080/02640414.2019.1592801

Saint-Maurice, P. F., Bai, Y., Vazou, S., & Welk, G. (2018). Youth physical activity patterns during school and out-of-school time. *Children*, *5*(9). https://doi.org/10.3390/children509011/8

Sasaki, J. E., John, D., & Freedson, P. S. (2011). Validation and comparison of ActiGraph activity monitors. *Journal of Science & Medicine in Sport*, *14*(5), 411–416. https://doi.org/10.1016/j.jsams.2011.04.003

SHAPE, America, MacDonald, L. C., Chepko, S., & Holt, S. A. (2014). *National standards & grade-level outcomes for K-12 physical education*. Human Kinetics.

Trost, S. G. (2001). Objective measurement of physical activity in youth: Current issues, future directions. *Exercise and Sport Sciences Reviews*, *29*(1), 32–36. https://doi.org/10.1097/00003677-200101000-00007

Trost, S. G., Mciver, K. L., & Pate, R. R. (2005). Conducting accelerometer-based activity assessments in field-based research. *Medicine and Science in Sports and Exercise*, *37*(11), S531. https://doi.org/10.1249/01.mss.0000185657.86065.98

Udermann, B. E., Mayer, J. M., Murray, S. R., & Sagendorf, K. (2004). Influence of cup stacking on hand-eye coordination and reaction time of second-grade students. *Perceptual and Motor Skills*, *98*(2), 409–414. https://doi.org/10.2466/pms.98.2.409-414

Utter, A. C., Robertson, R. J., Nieman, D. C., & Kang, J. (2002). Children’s OMNI scale of perceived exertion: Walking/running evaluation. *Medicine and Science in Sports and Exercise*, *34*(1), 139–144. https://doi.org/10.1097/00005768-200201000-00022