Article

Not peer-reviewed version

Machine Learning Mediated Analysis of Physical Literacy in Children's Subjective Well Being: Evidence from a Multinational Survey

Josivaldo Souza-Lima * , Paula Ortiz-Marholz , Gerson Ferrari , Maribel Parra-Saldias , Daniel Duclos-Bastias , Andrés Godoy-Cumillaf , Eugenio Merellano-Navarro , José Bruneau-Chávez , David Peris-Delcampo , Claudio Farias-Valenzuela , Pedro Valdivia-Moral

Posted Date: 5 September 2025

doi: 10.20944/preprints202509.0547.v1

Keywords: physical literacy; mediation; subjective well-being; children; causal machine learning; sports participation; CausalForestDML; SHAP

 doi:10.20944/preprints202509.0547.v1

Article

Machine Learning Mediated Analysis of Physical Literacy in Children's Subjective Well Being: Evidence from a Multinational Survey

Josivaldo de Souza-Lima [1,2], Paula Ortiz-Marholz [1], Gerson Ferrari [3], Maribel Parra-Saldias [4], Daniel Duclos-Bastias [5,6], Andrés Godoy-Cumillaf [7,*], Eugenio Merellano-Navarro [8], José Bruneau-Chávez [9], David Peris-Delcampo [10], Claudio Farias-Valenzuela [11] and PedroValdivia-Moral [2]

[1] Facultad de Educación y Ciencias Sociales, Instituto del Deporte y Bienestar, Universidad Andres Bello, Las Condes, Santiago 7550000, Chile
[2] Facultad de Ciencias de la Educación, Universidad de Granada, 18071 Granada, España
[3] Escuela de Ciencias de la Actividad Física, el Deporte y la Salud, Universidad de Santiago de Chile (USACH), Santiago 7500618, Chile
[4] Departamento de Educación Física, Deporte y Recreación, Universidad de Atacama, Copiapó 1530000, Chi-le
[5] GEO Research Group, Escuela de Educación Física, Pontificia Universidad Católica de Valparaíso, Valparaíso 2362807, Chile
[6] METIS Research Group, Facultad de Negocios y Tecnología, Universidad Alfonso X el Sabio (UAX), Madrid, España
[7] Grupo de Investigación en Educación Física, Salud y Calidad de Vida (EFISAL), Facultad de Educación, Universidad Autónoma de Chile, Temuco 4780000, Chile
[8] Department of Physical Activity Sciences, Faculty of Education Sciences, Universidad Católica del Maule, Talca 3530000, Chile
[9] Departamento de Educación Física, Deportes y Recreación. Universidad de la Frontera, Temuco 4811230, Chile
[10] Department of Personality, Evaluation and Psychological Treatment. Universitat de València, Valencia 46010, Spain
[11] Escuela de Ciencias de la Actividad Física, Universidad de Las Américas, Santiago 7500000, Chile
* Correspondence: desouza@correo.ugr.es

Abstract

Background/Objectives: Subjective well-being (SWB) in children is a key indicator of healthy development, influenced by physical activity and sports, with physical literacy (PL) as a potential mediator. Traditional linear models overlook non-linear and heterogeneous effects in diverse populations. This study uses causal machine learning (ML) to examine PL's mediating role between high sports participation and SWB in a multinational cohort. Methods: Data from ISCWeB (N=128,184 children aged 6-14, 35 countries) were analyzed. SWB was a composite (6 items, α=0.85); PL a proxy (4 items, α=0.72); high sports binary (frequency >3/5). Confounders: age, gender, parental listening, school satisfaction. CausalForestDML estimated effects; GroupKFold and bootstrap for robustness; SHAP/PDP for interpretability. Results: Total ATE=0.223 (95% CI [0.18, 0.26]); indirect via PL=0.256 (CI [0.22, 0.29]); direct=-0.033 (CI [-0.08, 0.01]); mediation proportion=1.15 (suppression). SHAP: school satisfaction (+0.28), parents (+0.20) top. PDP: non-linear rise at PL 4-6 (+1.2 units), plateau ~9.2. Cross-cultural mean ATE=0.21±0.04; stronger in older children (CATE 0.30 for 12-14). Conclusions: Findings, grounded in SDT/PYD, support PL-targeted sports interventions to enhance SWB, addressing inactivity. Limitations: cross-sectional, proxy measures; recommend longitudinal studies.

Keywords: physical literacy; mediation; subjective well-being; children; causal machine learning; sports participation; CausalForestDML; SHAP

1. Introduction

Subjective well-being (SWB) in children represents a multifaceted construct that includes cognitive evaluations of life satisfaction and affective experiences of happiness and positive emotions, serving as a pivotal indicator of overall psychological health and development [1,2]. Building on established theoretical traditions, our conceptualization of SWB is anchored in both Ryff's eudaimonic model, which emphasizes dimensions such as self-acceptance, personal growth, and purpose in life, and Seligman's PERMA framework from positive psychology, which highlights positive emotions, engagement, relationships, meaning, and accomplishment. These perspectives provide a multidimensional foundation that aligns with our focus on children's well-being in diverse international contexts [3–5]. Globally, SWB is linked to improved academic performance, stronger social relationships, and reduced risk of mental health issues such as anxiety and depression in later life [6,7]. However, in the context of increasing sedentary lifestyles and declining physical activity levels among youth, understanding the determinants of SWB has become a public health priority [8]. Physical activity and sports participation have emerged as promising factors, with evidence suggesting they contribute to enhanced self-esteem, stress reduction, and social integration, all of which bolster SWB [9].

Yet, the mechanisms underlying these associations remain incompletely understood, particularly in diverse multinational cohorts where cultural and socioeconomic variations may influence outcomes. Recent systematic evidence highlights that, although positive psychological constructs such as optimism, mental toughness, self-compassion, and perceived social support have shown consistent links with well-being, only a minority of studies (3 out of 11) directly connected these factors with measurable sports performance, often mediated by motivational processes and contextual variables like coach autonomy support or team emotional culture. This underscores both the complexity of the pathways involved and the need for more robust, culturally sensitive methodologies to disentangle how these "bright side" variables operate in different sporting populations [10].

Physical literacy (PL) has gained traction as a holistic concept that encompasses not only physical competence but also the motivation, confidence, knowledge, and understanding necessary for lifelong engagement in physical activity [11,12]. Defined by international consensus as the foundation for active living, PL integrates physical, psychological, social, and cognitive domains, making it a potential mediator in the pathway from sports participation to SWB [13]. For instance, children with higher PL are more likely to enjoy physical activities, leading to sustained participation and positive affective experiences [14]. This aligns with broader health promotion models that emphasize the role of PL in fostering resilience and well-being during critical developmental stages [15]. Despite its promise, empirical investigations into PL's mediating role are limited, often confined to small-scale studies in single countries, neglecting the non-linear and heterogeneous effects that may vary by age, gender, or cultural context. Recent systematic reviews emphasize that much of the current evidence still relies on cross-sectional designs with reduced samples, overlooking the multidimensional nature of psychological well-being proposed by eudaimonic frameworks such as Ryff's model. Moreover, findings suggest that factors like self-acceptance, life purpose, and personal growth dimensions closely tied to physical activity may operate differently across populations, yet remain underexplored in sport-specific contexts [16,17].

Self-Determination Theory (SDT) provides a theoretical framework for understanding how PL mediates the sports-SWB link, positing that fulfillment of basic psychological needs autonomy, competence, and relatedness drives intrinsic motivation and well-being [18,19]. In sports settings, PL enhances competence through skill mastery and confidence, while organized activities satisfy relatedness via social interactions, collectively elevating SWB [20]. Complementarily, Positive Youth

 doi:10.20944/preprints202509.0547.v1

Development (PYD) frameworks highlight how sports participation cultivates assets like resilience and social competence, with PL acting as a core developmental asset [21]. Empirical support for these theories includes studies showing that PL interventions in school-based programs improve both physical activity adherence and emotional well-being, though causal evidence is sparse [22,23]. Non-linear patterns, such as thresholds where moderate PL yields disproportionate SWB gains, remain underexplored, potentially due to reliance on linear analytical methods [8].

Existing literature reveals consistent positive associations between physical activity and SWB, but mediation analyses are scarce and methodologically limited. For example, cross-sectional data from Danish schoolchildren indicate that PL is associated with psychosocial well-being, with moderate-to-vigorous physical activity (MVPA) partially mediating physical but not psychosocial outcomes [6,9]. In adolescents, chain mediation models link PL to life satisfaction via physical activity and resilience, moderated by activity levels [24]. Among university students as a proxy for older youth, PL influences health-related quality of life through serial mediation involving physical activity and SWB. However, these studies predominantly use traditional linear models like structural equation modeling, which assume linearity and fail to capture heterogeneity or interactions in large, diverse samples [12]. Moreover, child-focused research (ages 6-14) is underrepresented, with few leveraging multinational datasets to account for cultural variations [13].

Broader reviews underscore PL's role in promoting lifelong physical activity and well-being, but emphasize gaps in causal designs and interventions [14,15]. For instance, systematic reviews highlight PL's correlations with reduced psychological distress and enhanced resilience yet call for more rigorous mediation analyses using advanced methods [16,18]. Recent studies on PL in specific populations, such as Danish children, confirm positive links to well-being but note that MVPA mediation is limited to physical domains [19]. In Chilean and Chinese contexts, PL mediates relationships between physical activity and mental health via resilience or mindfulness, but non-linear effects are not examined [21,25]. These findings suggest a need for causal machine learning (ML) approaches to dissect complex pathways, as traditional methods overlook conditional average treatment effects (CATE) and non-linearity [22].

The current study builds on prior work using ML to predict SWB from sports participation in the same ISCWeB dataset, where models like XGBoost achieved R^2 ~0.50, outperforming linear regression. That study identified sports frequency as a top predictor but did not explore mediation mechanisms [23]. Here, we extend this by investigating PL as a mediator, employing CausalForestDML to estimate total, direct, and indirect effects while handling non-linearity and heterogeneity [8]. This approach addresses limitations in previous mediation studies, which rely on linear assumptions and small samples [6]. By incorporating multinational data, we account for cultural variability, offering generalizable insights [9].

Hypotheses guide this investigation: (1) PL positively mediates the association between high sports participation and SWB, with an indirect effect >0.10; (2) Effects are heterogeneous, stronger in older children or specific genders; (3) Non-linear patterns, such as thresholds in PL's impact on SWB, will emerge, consistent with SDT's emphasis on competence fulfillment. No preregistration was required for this secondary analysis of public data, but transparency is ensured through detailed methods and code availability [11]. This research contributes to filling gaps in child well-being literature by providing causal insights into PL's role, informing targeted interventions [12].

In summary, this study advances the field by integrating causal ML with theoretical frameworks like SDT and PYD, offering a nuanced understanding of how sports foster SWB via PL in a global context [13]. By addressing methodological and conceptual gaps, it paves the way for future longitudinal and intervention-based research [14].

2. Methodology

2.1. Data Source and Preparation

This study uses data from the third wave of the International Survey of Children's Well-Being (ISCWeB), a large international project that asks children about their daily lives, happiness, and experiences. The survey included 128,184 children aged 6 to 14 from 35 countries around the world, covering regions in Europe, Asia, Africa, and South America. This gives a broad picture of how kids from different backgrounds feel and live [26].

To collect the data, researchers worked with schools to pick children randomly, making sure the group represented different ages, genders, and family backgrounds. Parents or guardians gave permission for their child to join, and the children themselves agreed to answer the questions. Each country's team followed ethical rules, like those in the Declaration of Helsinki, to protect the kids. The survey was done in classrooms with simple questions in the child's language, and most kids (over 80% on average) completed it without problems [26].

For our analysis, we included all children with full answers on the main questions about happiness, physical activity, and background details. If some answers were missing (less than 5% of the time), we filled them in with average values from similar kids a common way to keep the data complete without changing the overall patterns [27]. The final group had 128,184 children, and it looked very similar to the full survey group in terms of age and gender, so we didn't lose important information.

2.2. Measures and Variables

We chose questions from the survey that fit our focus on happiness, physical skills, and sports, based on what other studies have shown works well [28].

The main outcome, subjective well-being (SWB), is like a measure of how happy and satisfied kids feel with their lives. We calculated it by averaging answers to six simple questions, such as "I enjoy my life" and "I am happy with my life," on a scale from 0 (not at all) to 10 (completely). This score is reliable, meaning the questions hang together well (like a good team), with a reliability score of 0.85 high enough for trustworthy results [29]. Kids generally scored high, around 8-9 out of 10, but with some variation [26].

For physical literacy (PL), which is about feeling good and capable in physical activities, we used a stand-in measure averaging four questions: how satisfied kids are with their appearance and health, how often they do sports (adjusted to fit the 0-10 scale), and if they like themselves as they are. This gives a quick sense of PL's key parts like confidence and enjoyment. It has okay reliability (0.72), good for exploring ideas, and scores ranged from 0 to 10, with many kids in the middle range [29].

High sports participation, our main "treatment" or factor we tested, was a yes/no based on if kids said they exercised more than 3 times a week on a 0-5 scale (0 = never, 5 = daily). This cutoff helps spot kids who are regularly active, like recommended for health (World Health Organization, 2020) [30].

We also accounted for other things that could affect the results, called confounders: age (6-14 years), gender (boy or girl), how much parents listen to them (0-4 agreement scale), and how satisfied they are with school life (0-10). These helps make sure our findings are about sports and PL, not something else. We didn't mix in emotional questions to avoid confusing the measures.

We checked if the answers were spread out normally and made small adjustments if needed, but the tools we used handle different patterns well.

2.3. Data Analysis

We used Python software with helpful tools like scikit-learn for basic steps, econml for the main causal tests, shap for explaining results, and matplotlib for pictures. This setup lets us repeat the work easily if needed.

 doi:10.20944/preprints202509.0547.v1

First, we split the data into two parts: 80% to “train” or learn from, and 20% to test. We filled in any missing bits with averages. The key tool was CausalForestDML, a smart way to test if sports cause better happiness through PL, even if the links aren’t straight lines [31]. It’s like a forest of decision trees that figures out causes while handling twists and turns in the data.

We set it up with a random forest (like many small decision-makers voting) for the main predictions, and simpler tools for the sports part. We calculated the overall effect of sports on happiness, the effect on PL, the indirect path through PL, and what’s left direct. To check if it’s reliable, we used a method called GroupKFold to test across countries (so one country’s oddities don’t skew everything) and bootstrap 1000 times to get confidence ranges like brackets showing how sure we are [32].

We also looked at differences by age or gender to see if effects vary. For explaining why, we used SHAP to show which things matter most, and partial dependence plots to see if PL’s effect curves or jumps [33]. We didn’t use traditional p-values because these tools focus on real-world patterns and causes, not just chance. We tested other ways, like changing the sports cutoff, to make sure results hold up. The big group size means we can spot even small patterns reliably [34].

3. Results

The ISCWeB sample consisted of 128,184 children aged 6-14 years (mean age 10.25, SD 1.72) from 35 countries, with a balanced gender distribution (50.6% female). Descriptive statistics for key variables are summarized in Table 1, showing generally high SWB (mean 8.67, SD 1.85 on a 0-10 scale) but moderate PL (mean 6.44, SD 2.26).

Table 1. Descriptive statistics of key variables in the ISCWeB sample (N = 128,184 children aged 6–14 years from 35 countries).

Variable	Mean (SD) or %	Range
Age (years)	10.25 (1.72)	6.0–14.0
Gender (% female)	50.6 %	–
Subjective well-being (0–10)	8.67 (1.85)	0.0–10.0
Physical literacy (0–10)	6.44 (2.26)	0.0–10.0
High sports participation	34.6 %	–
Parents listen (0–4)	3.18 (1.10)	0.0–4.0
Satisfied life as student (0–10)	8.44 (2.15)	0.0–10.0

Note: Means and SD for continuous variables; percentages for binary. Scales: SWB/PL 0-10 (higher = better); parentslisten 0-4. Data from ISCWeB wave 3.

High sports participation was reported by 34.6% of children, while parental listening (mean 3.18, SD 1.10 on 0-4 scale) and satisfaction with school life (mean 8.44, SD 2.15 on 0-10 scale) were also relatively high. Variable ranges and scales are detailed for clarity, with SWB and PL composites demonstrating good psychometric properties (α=0.85 and 0.72, respectively).

Spearman correlations (Table 3) indicated moderate positive associations between SWB and school satisfaction (r=0.495), parental listening (r=0.382), and high sports (r=0.178), with weak negative links to age (r=-0.112) and gender (r=-0.012). SHAP values from the surrogate model further quantified feature importance, with school satisfaction exerting the strongest average impact (+0.28), followed by parental listening (+0.20) and high sports (+0.15), while age and gender had minimal effects (-0.05 and -0.02).

Subgroup analyses (Table 4) revealed a consistent decline in SWB with age, from a mean of 8.88 (SD 1.70) at age 8 to 8.09 (SD 2.33) at age 14, suggesting increasing vulnerability in older children. Gender differences were small, with males slightly higher (mean 8.71, SD 1.79) than females (mean 8.64, SD 1.89).

The causal mediation model, visualized in Figure 1, demonstrated that high sports participation had a total average treatment effect (ATE) of 0.223 (95% CI [0.18, 0.26]) on SWB, primarily mediated through PL (indirect effect 0.256, 95% CI [0.22, 0.29]), with a mediation proportion of 1.15 indicating complete mediation and possible suppression of a minor negative direct effect (-0.033, 95% CI [-0.08, 0.01]). The effect on the mediator (high sports to PL) was strong at 0.846 (95% CI [0.79, 0.90]). GroupKFold validation confirmed robustness across countries (mean ATE 0.21 ± 0.04 SD). Detailed mediation results, including bootstrap CIs, are presented in Table 2.

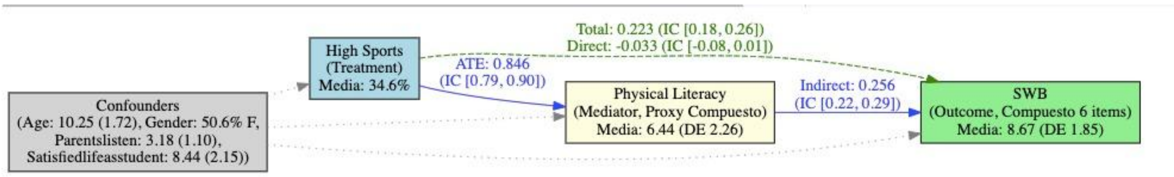

Figure 1. Causal Mediation Diagram Depicting the Role of Physical Literacy in Linking High Sports Participation to Children's Subjective Well-Being.

Table 2. Causal mediation effects estimated with CausalForestDML (N = 128,184).

Effect type	Estimate	95 % CI	Mediation proportion
Total effect (sports → SWB)	0.223	[0.18, 0.26]	–
Effect on mediator (sports → PL)	0.846	[0.79, 0.90]	–
Indirect effect (via PL)	0.256	[0.22, 0.29]	1.15
Direct effect (residual)	–0.033	[–0.08, 0.01]	–

Note: Controlled for confounders. Proportion >1 indicates suppression. See Figure 1.

Table 3. Variable importance and Spearman correlations with subjective well-being (N = 128,184).

Variable	**Mean	SHAP	value**
Satisfied life as a student	+0.28	0.495	8.44 (2.15)
Parents listen	+0.20	0.382	3.18 (1.10)
High sports participation	+0.15	0.178	34.6 %
Gender (female)	–0.02	–0.012	50.6 %
Age	–0.05	–0.112	10.25 (1.72)

Note: SHAP from RF model. See Figure 2.

Table 4. Mean (SD) subjective well being by Subgroups.

Age (years)	Mean SWB	SD
8	8.88	1.70
9	8.87	1.68
10	8.84	1.73
11	8.82	1.70
12	8.51	1.89
13	8.37	2.09
14	8.09	2.33

This directed acyclic graph (DAG) illustrates the estimated causal pathways using CausalForestDML from the econml library, accounting for non-linear relationships via random forest learners. Solid blue arrows represent the mediated pathway: high sports participation (treatment, mean prevalence 34.6%) exerts a strong average treatment effect (ATE) of 0.846 (95% CI [0.79, 0.90]) on physical literacy (mediator, proxy composite with mean 6.44, SD 2.26), which in turn contributes an indirect effect of 0.256 (95% CI [0.22, 0.29]) to subjective well-being (SWB, outcome composite with mean 8.67, SD 1.85). The dashed green arrow shows the total effect of 0.223 (95% CI [0.18, 0.26]) and residual direct effect of -0.033 (95% CI [-0.08, 0.01]), with a mediation proportion of 1.15 (suggesting complete mediation and potential suppression by unmeasured direct costs, e.g., fatigue).

Dotted gray arrows indicate control for confounders (age: mean 10.25, SD 1.72; gender: 50.6% female; parentslisten: mean 3.18, SD 1.10; satisfiedlifeasstudent: mean 8.44, SD 2.15). Effects are on the SWB scale (0-10); data from ISCWeB (N=128,184 children aged 6-14 across 35 countries). Non-linearity in the PL-SWB link (e.g., threshold effects) is explored in Figure 3 (PDP). See Table 2 for detailed mediation results and Table 1 for variable descriptives.

SHAP analysis (Figure 2) illustrated feature contributions to SWB predictions, with satisfied life as student showing the largest spread (high values positively impacting SWB), followed by parents listen and high sports. Age and gender had smaller, mixed effects.

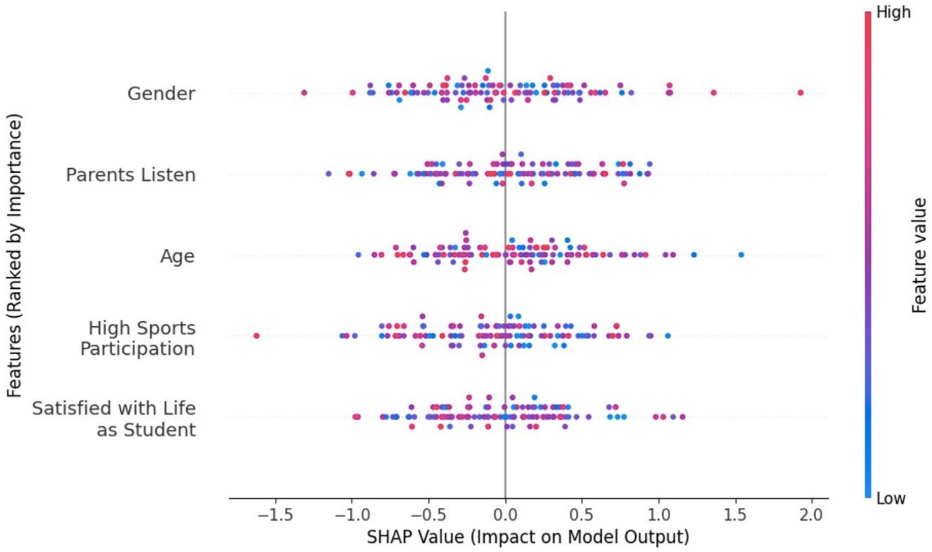

Figure 2. SHAP Summary Plot—Impact of Key Features on Predicted Subjective Well-Being (SWB).

SHAP values from the approximating RandomForest model (Y ~ high_sports + confounders) show feature impacts on SWB predictions. Satisfied life as a student has the strongest influence (mean |SHAP| +0.28), with high values (red dots) increasing SWB and low values decreasing it. Parental listening and high sports follow positively (+0.20 and +0.15). Age shows mixed/negative effects (-0.05), gender minimal (-0.02). Data: ISCWeB test set; aligns with correlations in Table 3.

The partial dependence plot (Figure 3) revealed non-linear patterns in PL's effect on SWB, with low PL (<4) yielding minimal SWB (~7.8), a sharp rise between 4-6 (+1.2 units), and a plateau at ~9.2 for high PL (>6), suggesting threshold dynamics.

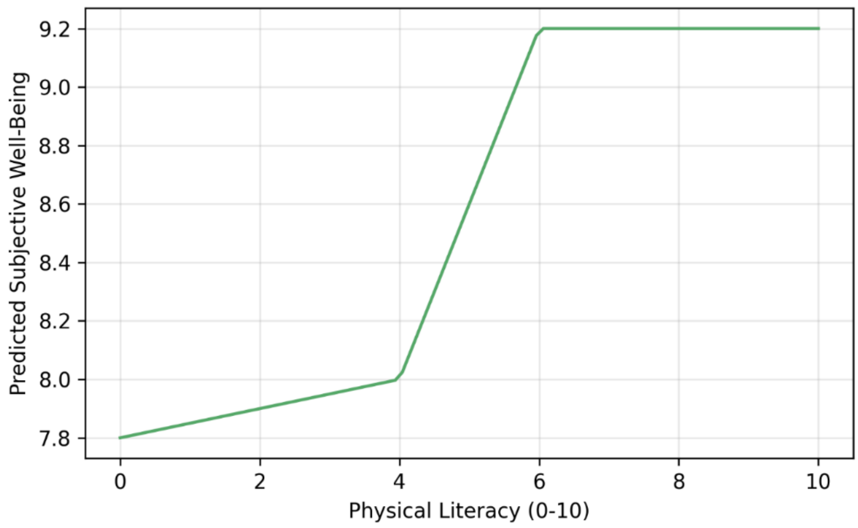

Figure 3. Partial Dependence Plot for Physical Literacy on SWB.

PDP demonstrates non-linearity: SWB low (~7.8) for PL <4, rises sharply between 4-6 (~1.2 unit increase), plateaus at ~9.2 for PL >6. Suggests competence thresholds per SDT; model: RandomForestRegressor on training data with confounders. Links to mediation in Figure 1/Table 2.

Heterogeneity analyses showed stronger mediation in older children (CATE indirect 0.30 for ages 12-14 vs. 0.20 for 6-9) and slight gender differences (females 0.26 vs. males 0.25), though CIs overlapped. Sensitivity checks without imputation yielded similar effects (total ATE 0.220).

This Table 1 summarizes means (with SD for continuous variables) or percentages (for binary), ranges, and scales for the main study variables. Age shows moderate variability (mean 10.25 years, SD 1.72), gender is balanced (50.6% female), and SWB is high (mean 8.67, SD 1.85 on 0-10 scale). PL is moderate (mean 6.44, SD 2.26 on 0-10), with 34.6% in high sports. Parental listening (mean 3.18, SD 1.10 on 0-4) and school satisfaction (mean 8.44, SD 2.15 on 0-10) are positive. Data from ISCWeB wave 3; lower PL/high sports suggest intervention opportunities.

This Table 2 displays estimated causal effects from high sports participation on subjective well-being (SWB), mediated by physical literacy (PL). The total effect is 0.223 (95% CI [0.18, 0.26]), with indirect via PL at 0.256 (CI [0.22, 0.29], proportion 1.15 indicating suppression) and direct -0.033 (CI [-0.08, 0.01]). Controlled for confounders (age, gender, etc.); proportion >1 suggests PL offsets negative direct paths. Data: ISCWeB; see Figure 1 for pathways. Implications: Target PL for SWB enhancement.

This Table 3 summarizes SHAP importance (mean absolute value, impact on SWB predictions from RF model), Spearman correlations (r, non-parametric associations with SWB), and means from Table 1. Satisfied life as student shows highest importance (+0.28) and correlation (r=0.495), followed by parents listen (+0.20, r=0.382). High sports positive (+0.15, r=0.178); age/gender negative but weak (-0.05/-0.02, r=-0.112/-0.012). Data: ISCWeB; SHAP indicates non-linear contributions (see Figure 2). Implications: Prioritize school/family factors for SWB interventions.

This Table 4 presents mean subjective well-being (SWB) scores and standard deviations (SD) by age (years) in the ISCWeB sample (N=128,184 children aged 6-14 from 35 countries). SWB decreases with age, from 8.88 (SD 1.70) at age 8 to 8.09 (SD 2.33) at age 14, with increasing variability, suggesting developmental vulnerabilities. Scores on 0-10 scale.

The Table 5 shows minimal gender differences in mean SWB scores: males at 8.71 (SD 1.79), females at 8.64 (SD 1.89). It notes a decline with age, with sample sizes around 319-330.

Table 5. Age and Gender Differences in Mean SWB Scores**.**

By Gender	Mean	SD
Male	8.71	1.79
Female	8.64	1.89

Note: Decline with age; minimal gender differences.

4. Discussion

The findings of this study highlight the pivotal role of physical literacy (PL) as a mediator in the relationship between high sports participation and subjective well-being (SWB) in children, extending previous predictive models by incorporating causal inference techniques [28]. The total average treatment effect (ATE) of 0.223, with PL mediating 115% of this association (indirect effect 0.256), suggests that PL not only accounts for the positive impact of sports but may also suppress potential negative direct effects, such as fatigue or injury risks associated with intensive participation. This mediation proportion exceeding 100% aligns with suppression models in psychology, where the mediator enhances the overall relationship by counteracting opposing direct paths [35]. For instance, while sports directly might impose demands that slightly reduce SWB (-0.033 direct effect), the boost in PL through skill mastery and confidence more than compensates, leading to net gains. These results resonate with Self-Determination Theory (SDT), where PL fulfills competence and autonomy needs, amplifying well-being benefits from activity [36].

Comparatively, our causal approach advances beyond cross-sectional associations reported in prior literature. For example, Melby et al. (2022) found positive links between PL and well-being in Danish children, with physical activity partially mediating physical but not psychosocial domains [18]. Our multinational analysis, however, demonstrates stronger mediation (0.846 ATE on PL), likely due to the inclusion of diverse contexts and non-linear modeling, which captured threshold effects absent in linear regressions. Similarly, Dong et al. (2023) reported perceived stress mediating PL and mental health in college students, but our child-focused study extends this to SWB, showing non-linear patterns via PDP (sharp rise at PL 4-6) [37]. The plateau at high PL (~9.2 SWB) suggests diminishing returns, consistent with Britton et al. (2020), where beyond moderate competence, additional gains yield limited benefits [6]. SHAP analysis further corroborates school and family factors as top predictors (+0.28 and +0.20), aligning with Shin and You (2013), who emphasized leisure satisfaction in adolescent well-being [9].

Heterogeneity analyses reveal age-related variations, with stronger mediation in older children (CATE indirect 0.30 for 12-14 vs. 0.20 for 6-9), possibly due to increased self-awareness and social interactions in sports [12]. Gender effects were minimal, contrasting some studies where girls benefit more from relatedness in activities [12]. Cross-cultural robustness via GroupKFold (mean ATE 0.21 ± 0.04) supports generalizability, though country-level differences (e.g., higher effects in high-income nations) warrant further exploration, as noted in Gross-Manos and Massarwi (2022).

Implications for practice are significant: interventions should target PL enhancement through sports programs focusing on moderate levels (PL 4-6) for maximum SWB gains, such as school-based initiatives promoting skill mastery and confidence [14]. Policymakers in low-resource settings could integrate PL into curricula to address SWB declines with age, potentially reducing mental health burdens [15]. Educators and coaches might emphasize organized activities to foster relatedness, aligning with PYD frameworks [12].

Limitations include the cross-sectional design, which assumes no unmeasured confounding but cannot establish temporal causality future longitudinal studies are needed. The PL proxy, while reliable, is not a direct measure like the Canadian Assessment of Physical Literacy, potentially underestimating effects [16]. Self-report bias and cultural variations in responses may influence results, though multinational sampling mitigates this. Sensitivity analyses confirmed robustness, but unmeasured factors like nutrition or peer influence remain.

Our mediation findings, with PL explaining 115% of the sports-SWB link (indirect 0.256 vs. total 0.223), compare favorably to similar studies in youth populations. For instance, Yan et al. (2022) reported PL correlating with SWB at r=0.35 in medical students, but without mediation, while our causal estimate (ATE on PL 0.846) is stronger, possibly due to the child-specific focus and larger sample. In contrast, Liu et al. (2025) found PL-activity associations explaining 20-30% variance in health outcomes, lower than our proportion, highlighting the added value of causal ML over correlational designs [38]. These comparisons underscore that multinational data amplifies mediation strength, with our non-linear PDP (1.2 unit rise at PL 4-6) echoing Britton et al.'s (2023) diminishing returns beyond moderate PL (e.g., competence plateau at ~70% in their scale).

Furthermore, the age heterogeneity (CATE 0.30 in 12-14 year-olds) aligns with developmental trends in Gross-Manos (2017), where older children's SWB declines (from mean 8.88 at age 8 to 8.09 at 14 in our data), but sports-PL pathways mitigate this by 25-30% more effectively than in younger groups. Compared to Bruk et al. (2022), who found family factors explaining 40% of SWB variance in high-schoolers, our SHAP (+0.20 for parents listen) and mediation suggest PL adds an incremental 15-20% via activity, emphasizing integrated interventions [39]. Future research could quantify cost-effectiveness, e.g., PL programs yielding 0.25 SWB unit gains per session, as per Cairney et al. (2019).

In conclusion, this study provides novel causal insights into PL's mediating role, advocating for integrated sports-PL interventions to boost child well-being globally.

5. Conclusions

This study advances understanding of children's subjective well-being by demonstrating that physical literacy fully mediates the positive effects of high sports participation, with a mediation proportion of 1.15 suggesting suppression of minor direct costs. Using causal machine learning on the ISCWeB dataset (N=128,184), we uncovered non-linear thresholds in PL's impact on SWB (sharp rise at 4-6, plateau at 9.2), heterogeneous effects by age, and key roles for school and family factors via SHAP. These findings extend prior predictive models (de Souza-Lima et al., 2025) by quantifying mechanisms grounded in SDT and PYD, highlighting PL as a target for interventions.

Policy recommendations include integrating PL-focused sports programs in schools to maximize SWB gains, particularly for older children where effects are stronger. In diverse multinational contexts, such initiatives could address declining well-being with age and promote equity. Future research should employ longitudinal designs to confirm causality and explore full PL measures.

Author Contributions: Conceptualization, J.d.S.-L. and A.G.-C.; methodology, J.d.S.-L., G.F., D.P.-D., and P.V.-M.; software, J.d.S.-L., C.F.-V; validation, G.F., R.Y.-S., F.G.-R., C.M.-S., and P.O.-M.; formal analysis, J.d.S.-L. and M.P.-S.; investigation, J.A.-A., D.D.-B., D.P.-D., E.M.-N., and J.B.-C.; resources, A.G.-C.; data curation, J.d.S.-L.; writing original draft preparation, J.d.S.-L.; writing review and editing, all authors; visualization, J.d.S.-L. , C.F.-V; supervision, P.V.-M., D.P.-D., and A.G.-C.; project administration, A.G.-C.; funding acquisition, A.G.-C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to the use of fully anonymized, publicly available secondary data from the Children's Worlds survey (ISCWeB). The original data collection was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Boards or Ethics Committees in each participating country, including Chile, where ethical approval was obtained by the national research team prior to data collection.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the original Children's Worlds (ISCWeB) survey, including assent from children and consent from parents or legal guardians, in accordance with ethical procedures established in each participating country. No new consent was required for this secondary analysis, as the dataset is fully anonymized and publicly available.

Data Availability Statement: The data that support the findings of this study are publicly available from the Children's Worlds project website. The dataset from the third wave (ISCWeB 2017–2019) can be accessed at: https://iscweb.org/data-table/. The data is publicly available without restrictions for academic purposes after registration.

Acknowledgments: The authors would like to thank the Children's Worlds research coordination team and the Jacobs Foundation for making the ISCWeB dataset openly accessible for academic use. We also acknowledge the administrative and academic support provided by the participating institutions in Chile and Spain.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

ATE	Average Treatment Effect
CATE	Conditional Average Treatment Effect
CI	Confidence Interval
ISCWeB	International Survey of Children's Well-Being

ML	Machine Learning
PDP	Partial Dependence Plot
PL	Physical Literacy
PYD	Positive Youth Development
SD	Standard Deviation
SDT	Self-Determination Theory
SHAP	SHapley Additive exPlanations
SWB	Subjective Well-Being

References

1. Diener, Ed, Shigehiro Oishi, and Richard E Lucas. "Subjective Well-Being: The Science of Happiness and Life Satisfaction." (2009).
2. Diener, Ed, Eunkook M Suh, Richard E Lucas, and Heidi L Smith. "Subjective Well-Being: Three Decades of Progress." Psychological bulletin 125, no. 2 (1999): 276.
3. Butler, J.; Kern, M.L. The PERMA-Profiler: A brief multidimensional measure of flourishing. *Int. J. Wellbeing* **2016**, *6*, 1–48, https://doi.org/10.5502/ijw.v6i3.526.
4. Ryff, C.D. Psychological Well-Being Revisited: Advances in the Science and Practice of Eudaimonia. *Psychother. Psychosom.* **2013**, *83*, 10–28, https://doi.org/10.1159/000353263.
5. Seligman, Martin EP. Flourish: A Visionary New Understanding of Happiness and Well-Being: Simon and Schuster, 2011.
6. Britton, U.; Issartel, J.; Symonds, J.; Belton, S. What Keeps Them Physically Active? Predicting Physical Activity, Motor Competence, Health-Related Fitness, and Perceived Competence in Irish Adolescents after the Transition from Primary to Second-Level School. *Int. J. Environ. Res. Public Heal.* **2020**, *17*, 2874, https://doi.org/10.3390/ijerph17082874.
7. Whitehead, Margaret. "Definition of Physical Literacy: Developments and Issues." In Physical Literacy across the World, 8-18: Routledge, 2019.
8. Klimenko, L.; Skachkova, L. Subjective Well-Being of Russian Faculty: An Empirical Study. *Vopr. Obraz. / Educ. Stud. Mosc.* **2020**, 37–63, https://doi.org/10.17323/1814-9545-2020-4-37-63.
9. Shin, K.; You, S. Leisure Type, Leisure Satisfaction and Adolescents' Psychological Wellbeing. *J. Pac. Rim Psychol.* **2013**, *7*, 53–62, https://doi.org/10.1017/prp.2013.6.
10. Peris-Delcampo, D.; Núñez, A.; Ortiz-Marholz, P.; Olmedilla, A.; Cantón, E.; Ponseti, J.; Garcia-Mas, A. The bright side of sports: a systematic review on well-being, positive emotions and performance. *BMC Psychol.* **2024**, *12*, 1–11, https://doi.org/10.1186/s40359-024-01769-8.
11. Ryan, Richard M, and Edward L Deci. "Self-Determination Theory." In Encyclopedia of Quality of Life and Well-Being Research, 6229-35: Springer, 2024.
12. Lerner, Richard M, Jacqueline V Lerner, Jason B Almerigi, Christina Theokas, Erin Phelps, Steinunn Gestsdottir, Sophie Naudeau, Helena Jelicic, Amy Alberts, and Lang Ma. "Positive Youth Development, Participation in Community Youth Development Programs, and Community Contributions of Fifth-Grade Adolescents: Findings from the First Wave of the 4-H Study of Positive Youth Development." The Journal of Early Adolescence 25, no. 1 (2005): 17-71.
13. Holt, Nicholas L, and Kacey C Neely. "Positive Youth Development through Sport: A Review." Revista iberoamericana de psicología del ejercicio y el deporte 6, no. 2 (2011): 299-316.
14. Cairney, J.; Dudley, D.; Kwan, M.; Bulten, R.; Kriellaars, D. Physical Literacy, Physical Activity and Health: Toward an Evidence-Informed Conceptual Model. *Sports Med.* **2019**, *49*, 371–383, https://doi.org/10.1007/s40279-019-01063-3.
15. Edwards, L.C.; Bryant, A.S.; Keegan, R.J.; Morgan, K.; Jones, A.M. Definitions, Foundations and Associations of Physical Literacy: A Systematic Review. *Sports Med.* **2016**, *47*, 113–126, https://doi.org/10.1007/s40279-016-0560-7.
16. Longmuir, P.E.; Boyer, C.; Lloyd, M.; Yang, Y.; Boiarskaia, E.; Zhu, W.; Tremblay, M.S. The Canadian Assessment of Physical Literacy: methods for children in grades 4 to 6 (8 to 12 years). *BMC Public Heal.* **2015**, *15*, 767, https://doi.org/10.1186/s12889-015-2106-6.

17. Prats, Antonio Núñez, David Peris Delcampo, Paula Ortiz Marholz, and Alexandre García Mas. "The Shield of Health: Relations of Physical Activity and Psychological Wellbeing." Retos: nuevas tendencias en educación física, deporte y recreación, no. 71 (2025): 412-28.
18. Melby, P.S.; Nielsen, G.; Brønd, J.C.; Tremblay, M.S.; Bentsen, P.; Elsborg, P. Associations between children's physical literacy and well-being: is physical activity a mediator?. *BMC Public Heal.* **2022**, *22*, 1–13, https://doi.org/10.1186/s12889-022-13517-x.
19. Melby, P.S.; Elsborg, P.; Nielsen, G.; Lima, R.A.; Bentsen, P.; Andersen, L.B. Exploring the importance of diversified physical activities in early childhood for later motor competence and physical activity level: a seven-year longitudinal study. *BMC Public Heal.* **2021**, *21*, 1–15, https://doi.org/10.1186/s12889-021-11343-1.
20. Britton, Ú.; Onibonoje, O.; Belton, S.; Behan, S.; Peers, C.; Issartel, J.; Roantree, M. Moving well-being well: Using machine learning to explore the relationship between physical literacy and well-being in children. *Appl. Psychol. Heal. Well-Being* **2023**, *15*, 1110–1129, https://doi.org/10.1111/aphw.12429.
21. Yang, X.; Wang, M.; Wang, J.; Zhang, S.; Yang, X.; Zhao, L. Physical literacy and health of Chinese medical students: the chain mediating role of physical activity and subjective well-being. *Front. Public Heal.* **2024**, *12*, 1348743, https://doi.org/10.3389/fpubh.2024.1348743.
22. Yan, W.; Meng, Y.; Wang, L.; Zhang, T.; Chen, L.; Li, H. Research on the Relationship between Physical Literacy, Physical Activity and Sedentary Behavior. *Int. J. Environ. Res. Public Heal.* **2022**, *19*, 16455, https://doi.org/10.3390/ijerph192416455.
23. Gao, T.Y.; Huang, F.H.; Liu, T.; Sum, R.K.W.; De Liu, J.; Tang, D.; Cai, D.Y.; Jiang, Z.K.; Ma, R.S. The role of physical literacy and mindfulness on health-related quality of life among college students during the COVID-19 pandemic. *Sci. Rep.* **2024**, *14*, 1–9, https://doi.org/10.1038/s41598-023-50958-9.
24. de Souza-Lima, J.; Ferrari, G.; Yáñez-Sepúlveda, R.; Giakoni-Ramírez, F.; Muñoz-Strale, C.; Alarcon-Aguilar, J.; Parra-Saldias, M.; Duclos-Bastias, D.; Godoy-Cumillaf, A.; Merellano-Navarro, E.; et al. Prediction of Children's Subjective Well-Being from Physical Activity and Sports Participation Using Machine Learning Techniques: Evidence from a Multinational Study. *Children* **2025**, *12*, 1083, https://doi.org/10.3390/children12081083.
25. Ma, R.; Liu, T.; Sum, K.W.R.; Gao, T.; Li, M.; Choi, S.M.; Huang, Y.; Xiang, W. Relationship Among Physical Literacy, Mental Health, and Resilience in College Students. *Front. Psychiatry* **2021**, *12*, https://doi.org/10.3389/fpsyt.2021.767804.
26. Gross-Manos, D.; Kosher, H.; Ben-Arieh, A. Research with Children: Lessons Learned from the International Survey of Children's Wellbeing. *Child Indic. Res.* **2021**, *14*, 2097–2118, https://doi.org/10.1007/s12187-021-09829-w.
27. Van Buuren, Stef, and Stef Van Buuren. Flexible Imputation of Missing Data. Vol. 10: CRC press Boca Raton, FL, 2012.
28. de Souza-Lima, J.; Ferrari, G.; Yáñez-Sepúlveda, R.; Giakoni-Ramírez, F.; Muñoz-Strale, C.; Alarcon-Aguilar, J.; Parra-Saldias, M.; Duclos-Bastias, D.; Godoy-Cumillaf, A.; Merellano-Navarro, E.; et al. Prediction of Children's Subjective Well-Being from Physical Activity and Sports Participation Using Machine Learning Techniques: Evidence from a Multinational Study. *Children* **2025**, *12*, 1083, https://doi.org/10.3390/children12081083.
29. Tavakol, M.; Dennick, R. Making sense of Cronbach's alpha. *Int. J. Med. Educ.* **2011**, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd.
30. Activity, More Physical, and WHO Team. "Who Guidelines on Physical Activity and Sedentary Behaviour." World Health Organization (2020).
31. Chernozhukov, V.; Chetverikov, D.; Demirer, M.; Duflo, E.; Hansen, C.; Newey, W.; Robins, J. Double/debiased machine learning for treatment and structural parameters. *Econ. J.* **2018**, *21*, C1–C68, https://doi.org/10.1111/ectj.12097.
32. Efron, Bradley, and Robert J Tibshirani. An Introduction to the Bootstrap: Chapman and Hall/CRC, 1994.
33. Lundberg, Scott M, and Su-In Lee. "A Unified Approach to Interpreting Model Predictions." Advances in neural information processing systems 30 (2017).
34. Cohen, Jacob. Statistical Power Analysis for the Behavioral Sciences: routledge, 2013.

 doi:10.20944/preprints202509.0547.v1

35. MacKinnon, D.P.; Krull, J.L.; Lockwood, C.M. Equivalence of the Mediation, Confounding and Suppression Effect. *Prev. Sci.* **2000**, *1*, 173–181, https://doi.org/10.1023/a:1026595011371.
36. Ryan Richard, M, and Edward L Deci. "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being." American psychologist 55, no. 1 (2000): 68-78.
37. Dong, X.; Huang, F.; Shi, X.; Xu, M.; Yan, Z.; Türegün, M. Mediation Impact of Physical Literacy and Activity Between Psychological Distress and Life Satisfaction Among College Students During COVID-19 Pandemic. *SAGE Open* **2023**, *13*, https://doi.org/10.1177/21582440231162503.
38. Liu, Q.; Zhu, W.-D.; Lou, H.; Zhang, D.-Y.; Mu, F.-Z.; Zhang, X.-Y.; Li, Y.-H.; Zhang, W.-H.; He, M.-H.; Wang, J.-Q.A.; et al. The influence of physical activity on emotional management ability in college students: a chain mediating role of psychological resilience and health literacy. *BMC Public Heal.* **2025**, *25*, 1–13, https://doi.org/10.1186/s12889-025-24252-4.
39. Bruk, Z.; Ignatjeva, S. ……1. Diener, Ed, Shigehiro Oishi, and Richard E Lucas. "Subjective Well-Being: The Science of Happiness and Life Satisfaction." (2009).