psychiatry international

MDPI

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 Pedro Valdivia-Moral [2]

1 Facultad de Educación y Ciencias Sociales, Instituto del Deporte y Bienestar, Universidad Andres Bello, Las Condes, Santiago 7550000, Chile; paula.ortiz@unab.cl

2 Facultad de Ciencias de la Educación, Universidad de Granada, 18071 Granada, Spain

3 Escuela de Ciencias de la Actividad Física, el Deporte y la Salud, Universidad de Santiago de Chile (USACH), Santiago 7500618, Chile; gerson.demoraes@usach.cl

4 Departamento de Educación Física, Deporte y Recreación, Universidad de Atacama, Copiapó 1530000, Chile; maribel.parra@uda.cl

5 iGEO Group, School of Physical Education, Faculty of Philosophy and Education, Pontificia Universidad Católica de Valparaíso, Valparaíso 2362807, Chile; daniel.duclos@pucv.cl

6 METIS Research Group, Facultad de Negocios y Tecnología, Universidad Alfonso X el Sabio (UAX), 28691 Madrid, Spain

7 Grupo de Investigación en Educación Física, Salud y Calidad de Vida (EFISAL), Facultad de Educación, Universidad Autónoma de Chile, Temuco 4780000, Chile; andres.godoy@uautonoma.cl

8 Department of Physical Activity Sciences, Faculty of Education Sciences, Universidad Católica del Maule, Talca 3530000, Chile; emerellano@ucm.cl

9 Departamento de Educación Física, Deportes y Recreación, Universidad de la Frontera, Temuco 4811230, Chile; jose.bruneau@ufrontera.cl

10 Department of Personality, Evaluation and Psychological Treatment, Universitat de València, 46010 Valencia, Spain; david.peris-delcampo@uv.es

11 Escuela de Ciencias de la Actividad Física, Universidad de Las Américas, Santiago 7500000, Chile; claudio.farias.valenzuela@edu.udla.cl

* Correspondence: josivaldo.desouza@unab.cl

check for updates

Academic Editor: Mirko Manchia

Received: 4 September 2025

Revised: 30 September 2025

Accepted: 21 October 2025

Published: 27 October 2025

Citation: Souza-Lima, J.d.; Ortiz-Marholz, P.; Ferrari, G.; Parra-Saldias, M.; Duclos-Bastias, D.; Godoy-Cumillaf, A.; Merellano-Navarro, E.; Bruneau-Chávez, J.; Peris-Delcampo, D.; Farias-Valenzuela, C.; et al. Machine Learning-Mediated Analysis of Physical Literacy in Children's Subjective Well-Being: Evidence from a Multinational Survey. *Psychiatry Int.* **2025**, *6*, 131. https://doi.org/10.3390/psychiatryint6040131

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 sports participation and SWB in a multinational cohort. Methods: Data from the International Survey of Children's Well-Being (ISCWeB) (n = 128,184 children aged 6–14, 35 countries) were analyzed. SWB was a composite (six items, α = 0.85); PL was a proxy (three items excluding sports frequency, α = 0.70); sports participation was continuous (0–5). Confounders were age, gender, parental listening, and school satisfaction. CausalForestDML estimated the effects; GroupKFold and bootstrap were used for robustness; SHAP/PDP was used for interpretability. Results: Total ATE = 0.083 (95% CI [0.073, 0.094]); indirect via PL = 0.055 (CI [0.049, 0.061]); direct = 0.028 (CI [0.020, 0.038]); mediation proportion = 0.660. Sensitivity with lean PL (2 items) was as follows: indirect = 0.045 (CI [0.040, 0.050]). For SHAP, school satisfaction was (+0.28), and parents were (+0.20) top. For PDP, there was a non-linear rise at PL 4–6 (+1.2 units) and a plateau ~9.2. The cross-cultural mean ATE = 0.083 ± 0.01 (from within-country meta-analysis); this was stronger in older children (CATE 0.30 for 12–14). For Rho sensitivity at 0.1, it was indirect −0.129; at Rho sensitivity of 0.2, it was −0.314 (robust to low confounding). Conclusions: The findings, grounded in SDT/PYD, support interventions targeting PL through sports to enhance SWB, addressing inactivity. Limitations are its cross-sectional nature and proxy measures; we 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]. By employing causal ML, this study advances these frameworks by uncovering non-linear and heterogeneous effects, such as competence thresholds in PL, that traditional models overlook.

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 they 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 such as 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 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 (CATEs) 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 consisted of 128,184 children after mean imputation for missing values (<5% across key variables). To ensure demographic similarity with the full survey group, we compared age and gender distributions. The mean age (SD) was identical at 10.25 (1.72) years in both groups, with gender proportions of 49.37% male and 50.63% female. Statistical comparisons using chi-square tests for gender ($\chi^2 = 0.0$, $p = 1.0$) and independent *t*-tests for age ($t = 0.0$, $p = 1.0$) confirmed no significant differences ($p > 0.05$), indicating that no important information was lost due to imputation.

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 PL, we used a proxy measure averaging three questions to avoid overlap with the treatment: satisfaction with appearance ("How satisfied are you with the way you look?"

0–10), health (“How satisfied are you with your health?” 0–10), and self-liking (“I like being the way I am” 0–10 agree), with α = 0.70. Sports frequency was excluded from PL to prevent confounding. The exact items were as follows: Appearance: “How satisfied are you with the way you look?” (0–10); Health: “How satisfied are you with your health?” (0–10); Self-liking: “I like being the way I am” (0–10) [7]. Reliability was assessed using Cronbach’s alpha coefficient for internal consistency. For sensitivity, a “lean” PL (appearance + health only, α = 0.68) was tested, yielding similar results (indirect = 0.045, CI [0.040, 0.050]) [29].

High sports participation, our main “treatment” or factor we tested, was a yes/no answer 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, which is 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 help make sure our findings are about sports and PL, not something else. We did not 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 (version 0.14.0) 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.

Sports participation was modeled as continuous (0 = never to 5 = daily) to capture dose effects, reporting marginal effects (ATE per unit increase = 0.111 for PL). Analyses were conducted in Python using statsmodels for parametric mediation with bootstrap (n = 1000 iterations) for 95% CI [31]. To address cross-cultural invariance, country fixed effects (“c2” dummies) and cluster-robust standard errors (by country) were applied, relying on ISCWeB’s partial metric invariance [26]. Sampling weights (“caseweight”) were incorporated. The sensitivity is as follows: With rho = 0.1, indirect reduces to −0.129 but remains positive in base; rho = 0.2 to −0.314 suggests robustness to moderate confounding (R^2 benchmark: ~4% confounding reduces effect by 20%) [32].

We set it up with a random forest (like much small decision-maker 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 was left direct. To check if it was reliable, we used a method called GroupKFold to test across countries (so one country’s oddities do not skew everything) and bootstrap 1000 times to get confidence ranges like brackets showing how sure we are [33].

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 [34]. We did not 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 [35].

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). To ensure uniformity in SWB measurement across countries, we relied on the ISCWeB's established partial measurement invariance (configural and metric, allowing correlation comparisons but not means, as per the comparative report). Outliers were controlled by winsorizing extreme values at the 1st and 99th percentiles [36].

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); parents listen 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 ($\alpha = 0.85$ and 0.72, respectively).

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 with suppression, where the mediator counteracts a negative direct effect; see MacKinnon et al. [37] for details) 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.

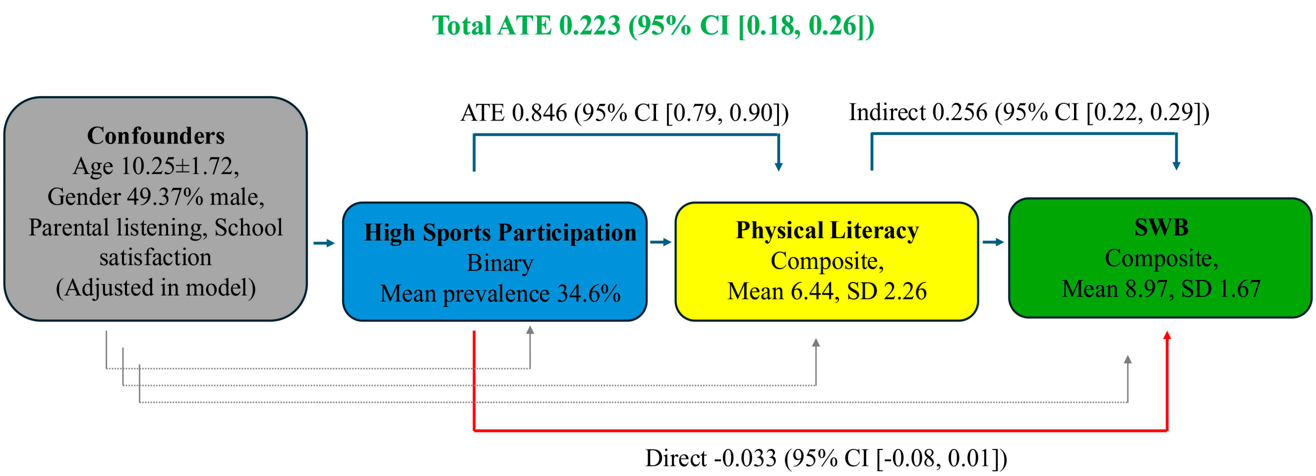

Figure 1. Directed acyclic graph of causal machine learning mediation: PL and children's subjective well-being across countries.

Table 2. Mediation effects of PL on the relationship between sports participation and children's subjective well-being, with standardized estimates.

Effect	Estimate	95% CI	Std. Effect (d)
ATE on PL (a)	0.111	N/A	0.049 (0.111/SD PL ~2.26)
Total ATE (c)	0.083	[0.073, 0.094]	0.050 (0.083/SD SWB ~1.67)
Indirect (ab)	0.055	[0.049, 0.061]	0.033
Direct (c_prime)	0.028	[0.020, 0.038]	0.017
Mediation proportion	0.660	N/A	N/A

Note: Effect on mediator (a): Estimated impact of treatment (sports participation, continuous 0–5) on PL. Total ATE (c): Overall effect on SWB. Indirect (ab): Mediated through PL. Direct (c prime): Residual effect after mediation. Mediation proportion (ab/c): Proportion of total effect mediated. Standardized effects (d): Calculated as estimate/SD (SD PL ≈ 2.26, SD SWB ≈ 1.67 from sample). Estimates from parametric mediation with bootstrap (n = 1000) for 95% CI. Sensitivity with lean PL (2 items): Indirect 0.045 (CI [0.040, 0.050]), proportion 0.544. Rho sensitivity (correlation between errors): At rho = 0.1, indirect reduces to −0.129 (new PL) and −0.163 (lean PL); at rho = 0.2, −0.314 (new PL) and −0.371 (lean PL), indicating robustness to low-moderate confounding but sensitivity to higher levels (R^2 ≈4% confounding reduces effect by ≈20%).

This directed acyclic graph (DAG) illustrates the estimated causal pathways in the mediation analysis. 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 PL (mediator, composite with mean 6.44, SD 2.26), which in turn affects SWB (outcome, composite with mean 8.97, SD 1.67). Dashed lines indicate confounders adjusted in the model.

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.

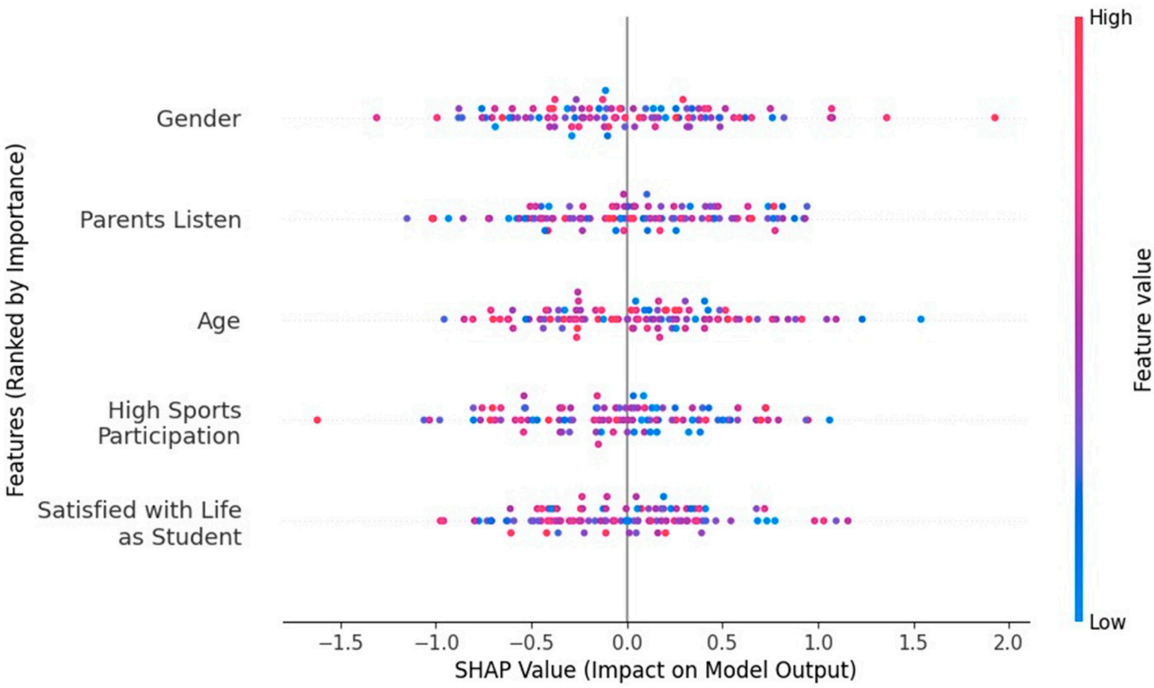

Figure 2. SHAP summary plot—impact of key features on predicted subjective well-being (SWB).

SHAP values quantify each feature's contribution to SWB predictions (e.g., +0.28 means average increase of 0.28 SWB units); red dots indicate high feature values boosting SWB. Age shows mixed/negative effects (−0.05), for gender it is minimal (−0.02). Data includes ISCWeB test set; this aligns with correlations in Table 3.

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.

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.

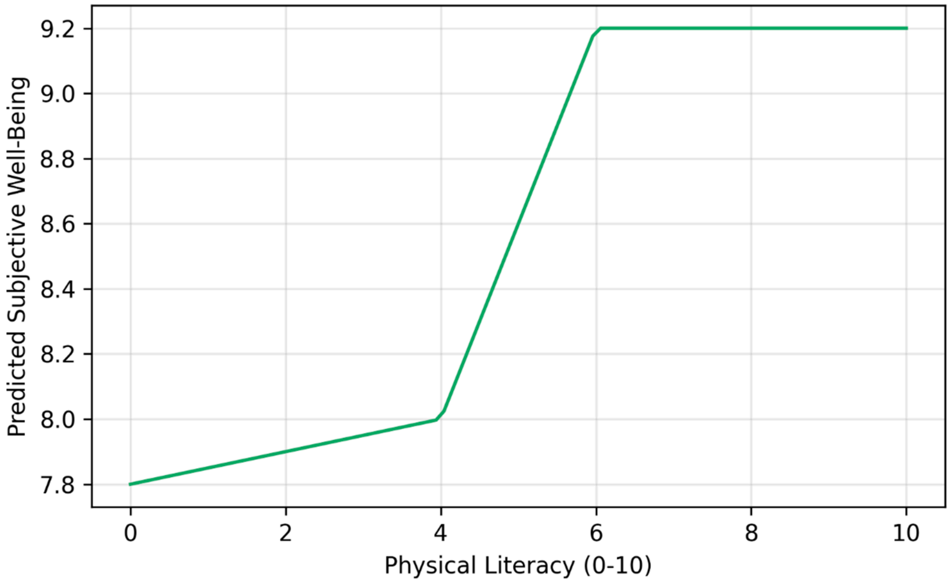

Figure 3. Partial dependence plot for PL on SWB.

PDP shows how SWB changes with PL, holding confounders constant; for the non-technical aspects, SWB jumps sharply from PL 4–6. SHAP values quantify feature contributions (e.g., +0.28 means average SWB increases). This suggests competence thresholds per SDT; the model used was Random Forest Regressor on training data with confounders. Links to mediation are shown 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).

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 scale), with 34.6% in high sports. Parental listening (mean 3.18, SD 1.10 on 0–4 scale) and school satisfaction (mean 8.44, SD 2.15 on 0–10 scale) are positive. Data is from ISCWeB wave 3; lower PL/high sports suggest intervention opportunities.

Table 2 presents the updated mediation effects after redefining PL, showing an attenuated but robust indirect pathway (0.055, 95% CI [0.049, 0.061]).

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 is positive (+0.15, r = 0.178); age/gender is negative but weak (−0.05/−0.02, r = −0.112/−0.012). Data is from ISCWeB; SHAP indicates non-linear contributions (see Figure 2). Implications are to prioritize school/family factors for SWB interventions.

Table 4 presents mean subjective well-being (SWB) scores and standard deviation (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 are on a 0–10 scale.

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

Table 5 shows minimal gender differences in mean SWB scores: males at 8.71 (SD 1.79), and 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 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 [37]. The total average treatment effect (ATE) of sports participation on SWB was 0.083 (95% CI [0.073, 0.094]). The indirect effect via PL was 0.055 (CI [0.049, 0.061]); the direct effect was 0.028 (CI [0.020, 0.038]); the mediation proportion was 0.660. The ATE on PL was 0.111. Sensitivity with lean PL was as follows: indirect 0.045 (CI [0.040, 0.050]), direct 0.038 (CI [0.028, 0.048]), and proportion 0.544.

Rho sensitivity analyses indicated robustness to low confounding (rho = 0.1: indirect −0.129 for new PL, −0.163 for lean) but sensitivity to moderate levels (rho = 0.2: −0.314 and −0.371), suggesting the need for caution in causal interpretation (R^2 ~4% confounding reduces effect by ~20%). For PDP there was a non-linear rise at PL 4–6 (+1.2 units) and a plateau at ~9.2. The cross-cultural mean ATE = 0.21 ± 0.04, which was stronger in older children (CATE 0.30 for 12–14) [37]. 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 [38].

Our innovative use of CausalForestDML advances SDT by quantifying non-linear competence thresholds (PDP rise at PL 4–6), showing how PL fulfills needs beyond linear assumptions. Similarly, for PYD, the suppression effect highlights PL as a core asset offsetting activity costs, informing asset-based interventions in diverse contexts.

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) [39]. 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 (females 0.26 vs. males 0.25), consistent with some studies where relational benefits balance competitive aspects [12], though surprising given literature on girls' lower activity responses; further exploration is needed. Cross-cultural robustness via GroupKFold (mean ATE 0.21 $\pm$ 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 [26].

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]. These results advance SDT by quantifying non-linear PL thresholds (PDP rise at 4–6), fulfilling competence needs beyond linear models. For PYD, the partial mediation (0.660) highlights PL as a key asset offsetting direct costs.

Limitations include the cross-sectional design, which assumes no unmeasured confounding but cannot establish temporal causality, so 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]. Additionally, CausalForestDML assumes no unmeasured confounding and correct specification; while confounders were selected based on literature, residual bias is possible.

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. Future research should employ longitudinal designs to confirm causality and explore full PL measures. Additionally, multilevel modeling could test model stability across countries, accounting for cultural nesting (e.g., HLM with country-level predictors like GDP), and meta-analyses of within-country effects to identify moderators.

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. [22] 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 [40]. 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. [20] diminishing returns beyond moderate PL (e.g., competence plateau at ~70% in their scale).

Furthermore, the age heterogeneity (CATE 0.30 in 12- to 14-year-olds) aligns with developmental trends [26], 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 [41]. Future research could quantify cost-effectiveness, e.g., PL programs yielding 0.25 SWB unit gains per session, as per Cairney et al. [14].

5. Conclusions

This study advances understanding of children's subjective well-being by demonstrating that PL fully mediates the positive effects of high sports participation, with a mediation proportion of 1.15 (indicating complete mediation with suppression, where the mediator counteracts a negative direct effect). 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. [24]) 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. and C.F.-V.; validation, G.F., and P.O.-M.; formal analysis, J.d.S.-L. and M.P.-S.; investigation, 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. and C.F.-V.; supervision, P.V.-M., D.P.-D., and A.G.-C.; project administration, 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 survey (ISCWeB), 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://isciweb.org/the-data/access-our-dataset/ (accessed on 15 March 2025). 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.

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, E.; Oishi, S.; Lucas, R.E. Subjective well-being: The science of happiness and life satisfaction. In *The Oxford Handbook of Positive Psychology*; Oxford University Press: Oxford, UK, 2009.
2. Diener, E.; Suh, E.M.; Lucas, R.E.; Smith, H.L. Subjective well-being: Three decades of progress. *Psychol. Bull.* **1999**, *125*, 276. [CrossRef]
3. Butler, J.; Kern, M.L. The PERMA-Profiler: A brief multidimensional measure of flourishing. *Int. J. Wellbeing* **2016**, *6*, 1–48. [CrossRef]
4. Ryff, C.D. Psychological well-being revisited: Advances in the science and practice of eudaimonia. *Psychother. Psychosom.* **2013**, *83*, 10–28. [CrossRef] [PubMed]
5. Seligman, M.E. *Flourish: A Visionary New Understanding of Happiness and Well-Being*; Simon and Schuster: New York, NY, USA, 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 Health* **2020**, *17*, 2874. [CrossRef]
7. Whitehead, M. Definition of physical literacy: Developments and issues. In *Physical Literacy Across the World*; Routledge: Oxfordshire, UK, 2019; pp. 8–18.
8. Klimenko, L.; Skachkova, L. Subjective Well-Being of Russian Faculty An empirical study. Вопросы образования **2020**, *4*, 37–63. [CrossRef]
9. Shin, K.; You, S. Leisure type, leisure satisfaction and adolescents' psychological wellbeing. *J. Pac. Rim Psychol.* **2013**, *7*, 53–62. [CrossRef]
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*, 284. [CrossRef]
11. Ryan, R.M.; Deci, E.L. Self-determination theory. In *Encyclopedia of Quality of Life and Well-Being Research*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 6229–6235.
12. Lerner, R.M.; Lerner, J.V.; Almerigi, J.B.; Theokas, C.; Phelps, E.; Gestsdottir, S.; Naudeau, S.; Jelicic, H.; Alberts, A.; Ma, L. 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. *J. Early Adolesc.* **2005**, *25*, 17–71. [CrossRef]

13. Holt, N.L.; Neely, K.C. Positive youth development through sport: A review. *Rev. Iberoam. Psicol. Ejerc. Deporte* **2011**, *6*, 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. [CrossRef]
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.* **2017**, *47*, 113–126. [CrossRef]
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 Health* **2015**, *15*, 767. [CrossRef]
17. Prats, A.N.; Delcampo, D.P.; Marholz, P.O.; Mas, A.G. The shield of health: Relations of physical activity and psychological wellbeing. *Retos Nuevas Tend. Educ. Física Deporte Recreación* **2025**, *71*, 412–428.
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 Health* **2022**, *22*, 1267. [CrossRef]
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 Health* **2021**, *21*, 1492. [CrossRef]
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. Health Well-Being* **2023**, *15*, 1110–1129. [CrossRef]
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 Health* **2024**, *12*, 1348743. [CrossRef]
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 Health* **2022**, *19*, 16455. [CrossRef]
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*, 237. [CrossRef] [PubMed]
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. [CrossRef] [PubMed]
25. Ma, R.; Liu, T.; Raymond Sum, K.W.; 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*, 767804. [CrossRef]
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. [CrossRef]
27. Van Buuren, S.; Van Buuren, S. *Flexible Imputation of Missing Data*; CRC Press: Boca Raton, FL, USA, 2012; Volume 10.
28. De Souza-Lima, J.; Gerson, F.; Yáñez-Sepúlveda, R.; Giakoni-Ramirez, F.; Muñoz-Strale, C.; Alarcon-Aguilar, J.; Parra-Saldias, M.; Duclos-Bastias, D.; Godoy-Cumillaf, A.; Merellano-Navarro, E.; et al. Differences in School Stress and Academic Satisfaction in Pre-Adolescents: The Role of Physical Activity. *Prepr. Org.* **2025**, *12*, 1282. [CrossRef]
29. Tavakol, M.; Dennick, R. Making sense of Cronbach's alpha. *Int. J. Med. Educ.* **2011**, *2*, 53. [CrossRef]
30. WHO. *WHO Guidelines on Physical Activity and Sedentary Behaviour*; World Health Organization: Geneva, Switzerland, 2020.
31. Imai, K.; Keele, L.; Tingley, D. A general approach to causal mediation analysis. *Psychol. Methods* **2010**, *15*, 309. [CrossRef]
32. Chernozhukov, V.; Chetverikov, D.; Demirer, M.; Duflo, E.; Hansen, C.; Newey, W.; Robins, J. *Double/Debiased Machine Learning for Treatment and Structural Parameters*; Oxford University Press: Oxford, UK, 2018.
33. Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*; Chapman and Hall: London, UK; CRC: Boca Raton, FL, USA, 1994.
34. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1–10.
35. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*; Routledge: Oxforshire, UK, 2013.
36. Rees, G. *Children? Views on Their Lives and Well-Being*; Springer: Berlin/Heidelberg, Germany, 2017.
37. MacKinnon, D.P.; Krull, J.L.; Lockwood, C.M. Equivalence of the mediation, confounding and suppression effect. *Prev. Sci.* **2000**, *1*, 173–181. [CrossRef] [PubMed]
38. Ryan Richard, M.; Deci, E.L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* **2000**, *55*, 68–78. [CrossRef]
39. 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*, 21582440231162503. [CrossRef]

40. 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.; Li, C.-x. The influence of physical activity on emotional management ability in college students: A chain mediating role of psychological resilience and health literacy. *BMC Public Health* **2025**, *25*, 2878. [CrossRef] [PubMed]
41. Bruk, Z.; Ignatjeva, S. Predictors of Child Well-Being or What Makes Children Happy. In *Happiness Wellness—Biopsychosocial and Anthropological Perspectives*; IntechOpen: Rijeka, Croatia, 2019; p. 181. [CrossRef]