EFFECT OF FOOTBALL'S SPORT EDUCATION MODEL AS INTERVENTION ON PHYSICAL FITNESS AND LEARNING ATTITUDE AMONG JUNIOR HIGH SCHOOL STUDENTS IN HEBEI CHINA

By

ZHANG JUNLONG

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

March 2025

FPP 2025 14

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

EFFECT OF FOOTBALL'S SPORT EDUCATION MODEL AS INTERVENTION ON PHYSICAL FITNESS AND LEARNING ATTITUDE AMONG JUNIOR HIGH SCHOOL STUDENTS IN HEBEI CHINA

By

ZHANG JUNLONG

March 2025

Chairman : Professor Soh Kim Geok, PhD
Faculty : Educational Studies

This study evaluates the effects of a 12-week Football Sport Education Model (SEM) on physical fitness and learning attitudes among junior high school students in Qinhuangdao City, Hebei, China. A total of 76 students (aged 13–15) from two randomly selected schools participated in a cluster-randomized controlled trial, with one school assigned to the SEM group and the other to the Standard Teaching (ST) group. Both groups engaged in football training three times a week (36 sessions, 40 minutes per session). Physical fitness was assessed using standardized tests from the Chinese Football Association, while learning attitudes were measured through the Students' Attitudes Toward Physical Education Scale (SPEA). The Generalized Estimating Equation (GEE) analysis revealed significant group-by-time interaction effects in agility (X^2=9.223, p=0.010), speed (X^2=6.901, p=0.032), lower body explosiveness (X^2=12.430, p=0.002), cardiovascular fitness (X^2=15.613, p=0.001), cognitive learning attitude

(X^2=25.872, p=0.001), affective learning attitude (X^2=20.281, p=0.001), and overall learning attitude (X^2=29.438, p=0.001), favoring the SEM group. Effect sizes indicated the most notable improvements occurred between weeks 6 and 12, particularly in agility (effect size = 0.895) and learning attitude (effect size = 2.375). No significant differences were found for BMI-for-age, flexibility, or upper body strength (p>0.05). These findings suggest that the SEM is more effective than standard teaching in enhancing both physical fitness and learning attitudes in football, with potential for broader implementation in junior high schools. Future research should explore optimal intervention duration and frequency to maximize the benefits of the SEM in physical education.

Keywords: Sport Education Model, football curriculum, physical fitness, learning attitude, Junior high school students

SDG: GOAL 4: Quality Education

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

KESAN MODEL PENDIDIKAN SUKAN BOLA SEPAK SEBAGAI INTERVENSI TERHADAP KECERGASAN FIZIKAL DAN SIKAP PEMBELAJARAN DALAM KALANGAN PELAJAR SEKOLAH MENENGAH RENDAH DI HEBEI, CHINA

Oleh

ZHANG JUNLONG

Mac 2025

Pengerusi : Profesor Soh Kim Geok, PhD
Fakulti : Pengajian Pendidikan

Kajian ini menilai kesan Model Pendidikan Sukan Bola Sepak (SEM) selama 12 minggu terhadap kecergasan fizikal dan sikap pembelajaran dalam kalangan pelajar sekolah menengah rendah di Kota Qinhuangdao, Hebei, China. Seramai 76 pelajar (berumur 13–15 tahun) dari dua sekolah yang dipilih secara rawak mengambil bahagian dalam kajian kawalan rawak berkelompok, dengan satu sekolah diperuntukkan kepada kumpulan SEM dan satu lagi kepada kumpulan Pengajaran Standard (ST). Kedua-dua kumpulan menjalani latihan bola sepak tiga kali seminggu (36 sesi, 40 minit setiap sesi). Kecergasan fizikal dinilai menggunakan ujian piawai yang disyorkan oleh Persatuan Bola Sepak China, manakala sikap pembelajaran diukur menggunakan Skala Sikap Pelajar Terhadap Pendidikan Jasmani (SPEA). Analisis Persamaan Anggaran Umum (GEE) menunjukkan kesan interaksi kumpulan-masa yang signifikan terhadap ketangkasan (X^2=9.223, p=0.010), kelajuan (X^2=6.901, p=0.032), keupayaan eksplosif bahagian

bawah badan (X^2=12.430, p=0.002), kecergasan kardiovaskular (X^2=15.613, p=0.001), sikap pembelajaran kognitif (X^2=25.872, p=0.001), sikap pembelajaran afektif (X^2=20.281, p=0.001), dan jumlah keseluruhan sikap pembelajaran (X^2=29.438, p=0.001), dengan kumpulan SEM menunjukkan peningkatan yang lebih ketara. Saiz kesan menunjukkan peningkatan paling ketara berlaku antara minggu ke-6 dan ke-12, terutamanya dalam ketangkasan (saiz kesan = 0.895) dan sikap pembelajaran (saiz kesan = 2.375). Tiada perbezaan yang signifikan ditemui untuk BMI, kelenturan, atau kekuatan bahagian atas badan (p>0.05). Dapatan kajian ini mencadangkan bahawa SEM lebih berkesan daripada pengajaran standard dalam meningkatkan kecergasan fizikal dan sikap pembelajaran dalam bola sepak, dengan potensi untuk diperluaskan ke sekolah menengah rendah lain. Kajian masa depan disyorkan untuk meneroka tempoh dan kekerapan intervensi yang optimum bagi memaksimumkan manfaat SEM dalam pendidikan jasmani.

Kata Kunci: Model Pendidikan Sukan, kurikulum bola sepak, kecergasan fizikal, sikap pembelajaran, pelajar sekolah menengah rendah

SDG: MATLAMAT 4: Pendidikan Berkualiti

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to Prof. Dr. Soh Kim Geok, my most sincere, compassionate, and patient professor and the chairperson of my supervisory committee. Her invaluable academic recognition and guidance enabled me to complete my PhD research in Malaysia. I am also profoundly thankful for her care and assistance in my daily life, which made me feel the warmth of a motherly figure in a foreign land. I admire Prof. Soh's rigorous academic attitude, keen scholarly insight, and patient yet constructive supervision. Her nurturing spirit, cheerful demeanour, and thoughtful care will forever remain in my heart.

I would also like to thank my supervisory committee member, Dr. Mohd Ashraff Mohd Anuar, for his invaluable advice, guidance, constructive feedback, and encouragement throughout my academic journey.

Next, I would like to express my sincere gratitude to my family. I sincerely appreciate their unwavering support for my decisions and infinite patience and understanding. It is because of them that I was able to complete my doctoral studies.

My gratitude also goes to my friends, who have supported me academically and personally. Thank you for being there for me.

Lastly, I want to thank myself. In moments of uncertainty, I made this most challenging yet meaningful decision. I am grateful for my choice, persistence, and efforts, which have once again proven my strength and determination.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy The members of the Supervisory Committee were as follows:

Soh Kim Geok, PhD
Professor
Faculty of Educational Studies
Universiti Putra Malaysia
(Chairman)

Mohd Ashraff bin Mohd Anuar, PhD
Senior Lecturer
Faculty of Educational Studies
Universiti Putra Malaysia
(Member)

ZALILAH MOHD SHARIFF, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:8 May 2025

TABLE OF CONTENTS

		Page
ABSTRACT		i
ABSTRAK		iii
ACKNOWLEDGEMENTS		v
APPROVAL		vii
DECLARATION		xi
LIST OF TABLES		xv
LIST OF FIGURES		xvii
LIST OF ABBREVIATIONS		xix
CHAPTER		
1	**INTRODUCTION**	1
	1.1 Background of the Study	1
	1.2 Problem Statement	4
	1.3 Research Objectives	6
	1.3.1 General Research Objectives	6
	1.3.2 Specific Research Objectives	6
	1.4 Research Hypothesis	7
	1.4.1 General Research Hypotheses	7
	1.4.2 Specific Research Hypotheses	7
	1.5 Significant of the Study	9
	1.5.1 Theoretical Significance	9
	1.5.2 Practical Significance	10
	1.6 Delimitations of the Study	11
	1.6.1 Subject Delimitation	11
	1.6.2 Measurement Delimitation	11
	1.6.3 Intervention Delimitation	12
	1.7 Research Limitation	12
	1.8 Conceptual and Operational Definitions of Terms	15
	1.8.1 Football	15
	1.8.2 Standard Teaching Approach	16
	1.8.3 Sport Education Model	17
	1.8.4 Junior High School Students	19
	1.8.5 Physical Fitness	19
	1.8.6 Physical Education Learning Attitude	20
2	**LITERATURE REVIEW**	22
	2.1 Football Players' Match Requirements and Characteristics	22
	2.1.1 Football Match Requirements	22
	2.1.2 Characteristics of Football Players	23
	2.2 Current Situation of Chinese Football Curriculum	24
	2.3 Definition of Terms	25
	2.3.1 Pedagogical Models	25
	2.3.2 Physical Fitness	26

2.3.3 Learning Attitude 28
2.3.4 Junior High School Students 29
2.4 Pedagogical Models in Physical Education 29
2.5 Effects of the Sport Education Model on Students' Physical Fitness 30
2.6 Effects of the Sport Education Model on Students' Learning Attitude 33
2.7 Other Factors Influencing Learning Attitude and Physical Fitness of Students 35
2.7.1 Family Income 36
2.7.2 Time Spent in Physical Activities during and Outside of School Hours 37
2.7.3 Growth and Nutritional Status 38
2.8 Theoretical Framework 39
2.8.1 Introduction of Sport Education Model 39
2.8.2 Embodied Cognitive Theory 42
2.8.3 Attitude Theory 45
2.9 Theoretical Framework Establishment 48
2.10 Conceptual Framework 50
2.10.1 Introduction to the FITT Principle 50
2.10.2 Physical Fitness 52
2.10.3 PE Learning Attitude 53

3 METHODOLOGY 56
3.1 Introduction 56
3.2 Research Design 56
3.3 Study Location 57
3.4 Population and Sampling Method 59
3.4.1 Target Population 59
3.4.2 Sampling Size Calculation 59
3.4.3 Criteria for Sampling 63
3.4.4 Sampling Method 63
3.5 Instruments 64
3.5.1 Socio-demographic Information 65
3.5.2 Physical Activity Questionnaire for Adolescents (PAQ-A) 65
3.5.3 Students' Attitudes toward Physical Education Scale (SPEA) 67
3.5.4 Assessment and Method for Physical Fitness 67
3.6 Validation of Questionnaires 68
3.6.1 Validity of the Questionnaire 69
3.6.2 Content of Validity of Physical Fitness Test 70
3.7 Intervention 71
3.7.1 Developing the Intervention Program 72
3.7.2 Intervention Teaching Program 73
3.7.3 Validation for Intervention Teaching Program 76
3.7.4 Control of Extraneous Variables 77
3.8 Pilot Study 81
3.8.1 Reliability for Questionnaires 82
3.8.2 Reliability of Physical Fitness Test 84

3.9 Research Permissions 85
3.10 Data Collection 85
3.11 Data Analysis 86

4 RESULT 90
4.1 Introduction 90
4.2 Preliminary Examination of Statistical Hypotheses 90
4.3 Descriptive Statistics of Demographic Variables and Dependent Variable at Baseline 91
4.4 Inferential Statistics 92
4.5 Effect of Sport Education Model on Physical Fitness 92
4.5.1 Body Composition 92
4.5.2 Flexibility 96
4.5.3 Agility 100
4.5.4 Speed 104
4.5.5 Upper Body Strength 108
4.5.6 Lower Body Explosiveness 113
4.5.7 Cardiovascular Fitness 117
4.6 Effect of Sport Education Model on Learning Attitude 121
4.6.1 Cognitive Learning Attitude 121
4.6.2 Affective Learning Attitude 125
4.6.3 Holistic Learning Attitude 129
4.7 Conclusion 133

5 DISCUSSION, CONCLUSION, IMPLICATION, AND RECOMMENDATION 135
5.1 Introduction 135
5.2 Discussion on Physical Fitness Variables 135
5.2.1 Discussion on Body Composition 135
5.2.2 Discussion on Flexibility 138
5.2.3 Discussion on Agility 141
5.2.4 Discussion on Speed 145
5.2.5 Discussion on Upper Body Strength 147
5.2.6 Discussion on Lower Body Explosiveness 149
5.2.7 Discussion on Cardiovascular Fitness 151
5.3 Discussion on Learning Attitude 152
5.3.1 Discussion on Cognitive Learning Attitude Dimension Score 153
5.3.2 Discussion on Affective Learning Attitude Dimension Score 154
5.3.3 Discussion on Holistic Learning Attitude Score 156
5.4 Conclusion 158
5.4.1 Physical Fitness 158
5.4.2 Learning Attitude 159
5.5 Research Implication 160
5.6 Recommendations for Future Studies 161

REFERENCES 165
APPENDICES 187
BIODATA OF STUDENT 236
LIST OF PUBLICATIONS 237

LIST OF TABLES

Table		Page
3.1	Effect Size Calculation for Research Variables (Mean and SD)	60
3.2	Physical Level Division	66
3.3	Physical Fitness Assessment Items and Methods	68
3.4	Personal Details of Questionnaire Experts	70
3.5	Personal Details of Questionnaire Experts	71
3.6	Summarised Intervention Teaching Program	75
3.7	Personal Details of Intervention Teaching Program Experts	76
3.8	Test-retest Reliability of the Socio-demographic Information (n=42)	83
3.9	Cronbach's Alpha Value of the Students' Learning Attitudes toward Physical Education Scale	83
3.10	Test-retest Reliability of the Students' Attitudes toward Physical Education Scale	84
3.11	Measurement Reliability of the Physical Fitness Tests	84
3.12	The Recruitment for Statistical Analysis towards Each Hypothesis	88
4.1	Descriptive Statistics (Mean and SE) of BMI-for-age for Control and Experimental Groups at Different Time Points	92
4.2	Result of GEE on BMI-for-age Score	93
4.3	Within-group Comparison of BMI-for-age at Different Time Points	94
4.4	Comparison of BMI-for-age between Groups at Different Time Points	95
4.5	Descriptive Statistics (Mean and SE) of Sit and Reach for Control and Experimental Groups at Different Time Points	96
4.6	Result of GEE on Sit and Reach Score	97
4.7	Analysis of Differences between Groups by Gender	97
4.8	Within-group Comparison of Sit and Reach at Different Time Points	98

4.9	Comparison of Sit and Reach between Groups at Different Time Points	99
4.10	Descriptive Statistics (Mean and SE) of Arrowhead Agility Test for Control and Experimental Groups at Different Time Points	100
4.11	Result of GEE on Arrowhead Agility Test Score	101
4.12	Analysis of Differences between Groups by Gender	101
4.13	Within-group Comparison of Arrowhead Agility Test at Different Time Points	102
4.14	Comparison of Arrowhead Agility Test between Groups at Different Time Points	103
4.15	Descriptive Statistics (Mean and SE) of 30-meter Sprint Test for Control and Experimental Groups at Different Time Points	104
4.16	Result of GEE on 30-meter Sprint Test Score	105
4.17	Analysis of Differences between Groups by Gender	105
4.18	Within-group Comparison of 30-meter Sprint Test at Different Time Points	106
4.19	Comparison of 30-meter Sprint Test between Groups at Different Time Points	107
4.20	Descriptive Statistics (Mean and SE) of Pull-up Test for Control and Experimental Groups at Different Time Points	108
4.21	Result of GEE on Pull-up Test Score	109
4.22	Analysis of Differences between Groups by Gender	110
4.23	Within-group Comparison of Pull-up Test at Different Time Points	111
4.24	Comparison of Pull-up Test between Groups at Different Time Points	112
4.25	Descriptive Statistics (Mean and SE) of Vertical Jump Test for Control and Experimental Groups at Different Time Points	113
4.26	Result of GEE on Vertical Jump Test Score	114
4.27	Analysis of Differences between Groups by Gender	114
4.28	Within-group Comparison of Vertical Jump Test at Different Time Points	115
4.29	Comparison of Vertical Jump Test between Groups at Different Time Points	116

4.30	Descriptive Statistics (Mean and SE) of YO-YO Test for NT and SEM Groups at Different Time Points	117
4.31	Result of GEE on YO-YO Test Score	118
4.32	Analysis of Differences between Groups by Gender	118
4.33	Within-group Comparison of YO-YO Test at Different Time Points	119
4.34	Comparison of YO-YO Test between Groups at Different Time Points	120
4.35	Descriptive Statistics (Mean and SE) of Cognitive Dimension Score for Control and Experimental Groups at Different Time Points	121
4.36	Result of GEE on Cognitive Dimension Score	122
4.37	Within-group Comparison of Cognitive Dimension Score at Different Time Points	123
4.38	Comparison of Cognitive Dimension Score between Groups at Different Time Points	124
4.39	Descriptive Statistics (Mean and SE) of Affective Dimension Score for Control and Experimental Groups at Different Time Points	125
4.40	Result of GEE on Affective Dimension Score	126
4.41	Within-group Comparison of Affective Dimension Score at Different Time Points	127
4.42	Comparison of Affective Dimension Score between Groups at Different Time Points	128
4.43	Descriptive Statistics (Mean and SE) of Holistic Learning Attitude Score for Control and Experimental Groups at Different Time Points	129
4.44	Result of GEE on Holistic Learning Attitude Score	130
4.45	Within-group Comparison of Holistic Learning Attitude Score at Different Time Points	131
4.46	Comparison of Holistic Learning Attitude Score between Groups at Different Time Points	132

LIST OF FIGURES

Figure		Page
2.1	The Characteristics of the Sport Education Model	42
2.2	Embodied Cognitive Theory	45
2.3	Attitude Theory	48
2.4	Theoretical Framework	50
2.5	Conceptual Framework	55
3.1	Research Flow	57
3.2	Sample Size Calculation Based on the Repeated Measure within-between Interaction	61
3.3	Participants' Flow Diagram	64
4.1	The Mean Scores Changes of BMI-for-age across the Times for Both Groups	96
4.2	The Mean Scores Changes of Flexibility across the Times for Both Groups	100
4.3	The Mean Scores Changes of Agility across the Times for Both Groups	104
4.4	The Mean Scores Changes of Speed across the Times for Both Groups	108
4.5	The Mean Scores Changes of Muscle Endurance across the Times for Both Groups	112
4.6	The Mean Scores Changes of Power across the Times for Both Groups	117
4.7	The Mean Scores Changes of Cardiovascular Fitness across the Times for Both Groups	121
4.8	The Mean Score Changes of Cognitive Learning Attitude across the Times for Both Groups	125
4.9	The Mean Score Changes of Affective Learning Attitude across the Times for Both Groups	129
4.10	The Mean Scores Changes of Holistic Learning Attitude across the Times for Both Groups	133

LIST OF ABBREVIATIONS

SEM	Sport Education Model
PE	Physical Education
FIFA	International Federation of Association Football
ST	Standard Teaching
RCTs	Randomised Controlled Trials
SPEA	Students' Attitudes Toward Physical Education Scale
BIA	Bioelectrical Impedance Analysis
DXA	Dual-energy X-ray Absorptiometry
UNESCO	The United Nations Educational, Scientific, and Cultural Organisation
TGfU	Teaching Games for Understanding
TPSR	Teaching Personal and Social Responsibility
HBPE	Health-Based Physical Education
BMI-for-age	Body Mass Index-for-age
PFI	Physical Fitness Index
CRCT	Cluster Randomised Controlled Trial
PAQ-A	Physical Activity Questionnaire for Adolescents
MVPA	Moderate to Vigorous Physical Activity

CHAPTER 1

INTRODUCTION

This chapter mainly describes the background of the study, problem statement, research objectives, hypothesis, significant of the study, delimitations, limitations, and conceptual and operational definitions of terms.

1.1 Background of the Study

Football, often regarded as the world's most popular sport, significantly promotes physical and mental development among youth (Ihle et al., 2021; Nishanbayevich, 2022). Football has garnered increasing attention in China due to its efforts to improve its international competitiveness, particularly at the youth level. The sport's accessibility and appeal make it a powerful tool for physical education programs (Abdullaeva, 2021).

Physical education (PE) plays a fundamental role in shaping students' physical fitness and learning attitudes, especially during their school years. PE aims to improve students' physical fitness by providing regular, structured physical activity. Through activities such as running, jumping, and strength exercises, students develop critical components of fitness, including endurance, flexibility, coordination, and strength (Tinôco et al., 2023). Improved physical fitness has been linked to better academic performance, as students with higher fitness levels demonstrate greater concentration and memory retention (World Health Organisation, 2021).

Beyond physical fitness, PE plays a crucial role in shaping students' learning attitudes, including cognitive and affective components (Zhang et al., 2024). On the cognitive side, physical education encourages critical thinking and problem-solving as students learn to plan, execute, and reflect on their performance in various physical tasks. These experiences develop decision-making skills and strategic thinking transferable to academic contexts (Mallen & Grenville-Cleave, 2017). Regarding affective learning, PE promotes positive emotional experiences by fostering teamwork, collaboration, and a sense of achievement (Rivera-Pérez et al., 2021; Liu & Lipowski, 2021). The inclusive and cooperative nature of physical activities helps students build self-confidence and social skills, which are critical to maintaining motivation and engagement in physical activities (Lubans et al., 2022; Beni et al., 2021). Through successful participation in physical education, students develop emotional resilience, which positively influences their attitudes toward learning.

PE benefits students across various stages of their academic journey, from elementary to high school. For younger students, PE helps lay the foundation for physical literacy, encouraging the development of motor skills and a basic understanding of health and fitness (Young et al., 2021; Melby et al., 2021). As students advance to higher grades, PE fosters more complex cognitive and social skills, such as teamwork, leadership, and decision-making, which are crucial during adolescence (Young et al., 2021; Neil-Sztramko, 2021). At the high school level, PE is essential in maintaining physical fitness while preparing students for lifelong healthy habits and engagement in physical

activities (Lubans et al., 2021; Young et al., 2021). The structured and progressive nature of PE throughout these stages ensures that students maintain their physical health and develop the cognitive and emotional competencies necessary for success in both academic and personal life.

In recent years, there has been increasing interest in examining how various teaching models can enhance both physical fitness and learning attitudes (Méndez-Giménez et al., 2022; Zhen-Rong, 2021; Zhang et al., 2023; Zhang et al., 2024). One such model is the Sport Education Model (SEM), which emphasises student-centred learning, teamwork, and role-based participation. SEM offers a more structured and engaging approach to PE, allowing students to take on roles such as team captain or referee, promoting physical activity and cognitive engagement (Bessa et al., 2019). While the potential of SEM will be explored further in this study, its inclusion highlights the growing interest in adopting innovative methods to improve PE outcomes.

However, it is crucial to consider the unique developmental characteristics of students aged 13 to 15 years when assessing the effectiveness of PE interventions. Adolescence is marked by significant physiological changes, including growth spurts, variations in fitness levels, and hormonal fluctuations, which may impact students' physical performance, motor skill development, and learning attitudes (Malina et al., 2004; Cumming et al., 2017). These changes often lead to increased disparities in strength, endurance, and coordination between students of the same age, which could influence their responsiveness to training interventions (Baquet et al., 2010).

Furthermore, cognitive and emotional development during this period affects students' motivation, self-perception, and engagement in sports activities (Piaget, 1972). Given these factors, it is essential to examine whether and how these age-related variations influence the effectiveness of the Sport Education Model, ensuring that it can be effectively applied within junior high school settings.

1.2 Problem Statement

Despite China's significant investment in youth football development, its performance in international competitions, such as the FIFA World Cup, remains underwhelming (Sun, 2022). This issue is particularly evident in Hebei Province, where students' motivation for football and participation rates declines sharply upon entering junior high school (Xue & Li, 2023).

A key factor contributing to this decline is the overly structured and teacher-centered nature of standard football curricula, which focus primarily on skill acquisition while limiting student autonomy and engagement (Zhang, 2024; Fang, 2024). The rigid assessment system, which emphasizes performance-based evaluations, further discourages students from active participation (Xu et al., 2021). As a result, many students lose interest in football, leading to decreased motivation and reduced physical activity levels.

Moreover, national fitness monitoring data reveal a continuous decline in the physical fitness of Chinese students over the past three decades, with significant deterioration in cardiovascular fitness, muscular strength, and

flexibility (Dong et al., 2019; Wang, 2024). The inadequacy of standard physical education (PE) methods in promoting long-term engagement further exacerbates this issue, highlighting the need for a more student-centered approach.

The SEM presents a promising alternative. Unlike traditional methods, SEM emphasizes student engagement, teamwork, and authentic sports experiences, fostering both physical fitness and positive learning attitudes (Bessa et al., 2021). The decision to select SEM over other pedagogical models stems from its comprehensive and inclusive structure, which aligns closely with the goals of improving physical fitness and fostering positive attitudes towards physical education in adolescents. Traditional models, such as the teacher-centered approach, often fail to engage students meaningfully, particularly in middle school, where motivation and self-regulation are still developing. In contrast, SEM's focus on student roles (players, coaches, referees) and the creation of long-term team dynamics helps to build a deeper sense of engagement and commitment to physical activity. However, research on SEM in China remains limited, particularly at the junior high school level. Existing studies have primarily focused on high school and university students, leaving a gap in understanding its effectiveness among younger adolescents (Zhang et al., 2023; Zhang et al., 2024).

Furthermore, prior studies on SEM have often lacked rigorous research designs, with many relying on quasi-experimental methods without control groups (Zhang et al., 2024). Additionally, most studies have examined either

physical fitness or learning attitudes in isolation, rather than exploring their combined effects (Hastie et al., 2009; Wahl-Alexander et al., 2018; Chu et al., 2022). Given these gaps, there is a pressing need for well-designed research to evaluate the impact of an SEM-based football curriculum on both physical fitness and learning attitudes among junior high school students in China.

1.3 Research Objectives

The research objectives of this study are divided into general and specific objectives.

1.3.1 General Research Objectives

This study evaluated the effects of the SEM on junior high school students' physical fitness and learning attitude in Hebei, China.

1.3.2 Specific Research Objectives

1) To examine the effects of the SEM across baseline, post-test1, and post-test2 on body composition, flexibility, agility, speed, upper body strength, lower body strength, and cardiovascular fitness among junior high school students.
2) To examine the effects of the SEM across baseline, post-test1, and post-test2 on cognitive learning attitude, affective learning attitude, and holistic learning attitude among junior high school students.

1.4 Research Hypothesis

1.4.1 General Research Hypotheses

The first hypothesis of this study is to evaluate whether there are significant differences between the intervention group (SEM) and the control group (ST) in terms of physical fitness (body composition, flexibility, agility, speed, upper body strength, lower body strength, and cardiovascular fitness) across baseline, post-test1, and post-test2 on junior middle school students in Hebei, China.

The second hypothesis of this study is to evaluate whether there are significant differences between the intervention group (SEM) and the control group (ST) in terms of learning attitude (cognitive, affective, and holistic learning attitude) across baseline, post-test1, and post-test2 on junior middle school students in Hebei, China.

1.4.2 Specific Research Hypotheses

H01-1: There were no statistically significant differences between the experimental and control groups in body composition at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.

H01-2: There were no statistically significant differences between the experimental and control groups in flexibility at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.

H01-3: There were no statistically significant differences between the experimental and control groups in agility at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.

H01-4: There were no statistically significant differences between the experimental and control groups in speed at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.

H01-5: There were no statistically significant differences between the experimental and control groups in Upper body strength at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.

H01-6: There were no statistically significant differences between the experimental and control groups in Lower body explosiveness at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.

H01-7: There were no statistically significant differences between the experimental and control groups in cardiovascular fitness at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.

H02-1: There were no statistically significant differences between the experimental and control groups in cognitive learning attitude at baseline,

post-test1, and post-test2 among junior middle school students in Hebei, China.

H02-2: There were no statistically significant differences between the experimental and control groups in affective learning attitudes at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.

H02-3: There were no statistically significant differences between the experimental and control groups in holistic learning attitude at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.

1.5 Significant of the Study

This study applied specific theoretical and practical significance in PE and pedagogy. The details are explained below.

1.5.1 Theoretical Significance

This study pioneers applying Siedentop's SEM (Siedentop et al., 2011) and Varela et al.'s Embodied Cognition Theory (Varela et al., 2017) to the specific football context. Integrating these two theoretical frameworks offers a novel approach to explaining how the contextualised teaching of football through the SEM can enhance students' physical fitness and learning attitudes. Although these theories have been widely applied in various sports and exercise settings, research on using the Embodied Cognition Theory to

explain the impact of SEM programs on improving students' physical fitness and learning attitudes remains limited. This experimental study aims to evaluate the effects of football teaching under the guidance of the SEM through a 12-week SEM-based intervention on supporting a dynamic systems theory model of physical fitness and learning attitudes among junior high school students. This finding related to the effects may reveal the processes and mechanisms through which football SEM programs enhance students' physical fitness and learning attitudes. The results of this study could provide valuable insights for junior high school students, helping them to develop customised football teaching programs that improve physical fitness and learning attitudes, thereby contributing positively to the development of the sport, particularly by addressing the decline in physical fitness and enhancing learning attitudes toward football among junior high school students.

1.5.2 Practical Significance

The practical significance of implementing a 12-week junior high school football curriculum using the SEM lies in its potential to benefit multiple stakeholders. For education departments, this program could inspire meaningful reforms in physical education, elevating the quality of PE programs and aligning with national educational objectives. The SEM provides a structured yet flexible model for teachers that enhances teaching strategies and deepens their understanding of student learning dynamics. Moreover, this research supports students' physical development, nurtures a passion for football, and fosters improved learning attitudes by creating an engaging, motivational environment. If proven successful, this program could

serve as a valuable reference for designing football SEM curricula across various educational levels, including elementary schools, high schools, and universities in China.

1.6 Delimitations of the Study

Aligned with the ongoing trend of sports teaching reform in junior high school, this study leverages insights from the attributes and benefits of the SEM in the United States. It formulates a curriculum for the SEM in junior high school football, tailored to the specific context of China, and conducts empirical research. The primary objective is to investigate the implementation of the SEM in junior high school physical education instruction and its impact on student's physical fitness and attitudes toward PE learning. The study has some delimitations in terms of:

1.6.1 Subject Delimitation

Due to potential differences in educational systems across countries, this study focuses on junior high school students in Hebei Province, China, implying that participants are aged between 13 and 15 (Ouyang, 2023; Yang, 2022). Therefore, this study only selected junior high school students aged 13 to 15, including boys and girls.

1.6.2 Measurement Delimitation

Although many physical fitness and learning attitude variables could be adopted and tested, this study selected only seven physical fitness variables

(body composition, flexibility, agility, speed, upper body strength, lower body strength, and cardiovascular fitness) and two learning attitude variables (cognitive and affective dimensions). These variables are closely related to previous research on physical fitness and learning attitudes. This study's standards for physical fitness testing were based on the "Guidelines for Testing Athletic Abilities of Chinese Youth Football Players," jointly issued by the Chinese Football Association and the General Administration of Sport of China. The measurement tool for learning attitudes, the Students' Attitudes Toward Physical Education Scale (SPEA), is widely recognised as an effective instrument for accurately measuring learning attitudes (Subramaniam & Silverman, 2000; Hu et al., 2014).

1.6.3 Intervention Delimitation

Compared to standard teaching methods, the SEM offers a highly suitable approach for enhancing junior high school students' physical fitness and learning attitudes. This study developed a football teaching plan tailored to junior high school students based on the recommended content for Level Four in The Curriculum Standards for Physical Education and Health in Compulsory Education (2022 edition). The plan was carefully designed with expert input to ensure its suitability and effectiveness for this age group, making it an appropriate intervention for the targeted population.

1.7 Research Limitation

The intervention in this study follows the newly issued National Curriculum Standards (2022 edition) for junior high students (Level 4), designed for the

national average skill level. However, the actual skill level of students in Hebei Province may differ, potentially causing learning to be either too difficult or too easy, thereby affecting learning attitudes and emotions. Nonetheless, national standards provide advantages, as the content has been scientifically validated and fits the students' context. Additionally, many studies have adopted these standards with reliable results (Ren, 2024; Liang, 2024; Wang et al., 2023). In this study, the intervention protocol was also reviewed by six experts, further enhancing its validity.

Regarding physical fitness measurement, BMI-for-age was used to assess body composition. While BMI-for-age is widely utilised, research suggests more accurate methods, such as bioelectrical impedance analysis (BIA) and dual-energy X-ray absorptiometry (DXA), are preferable for assessing body fat percentage (Bolanowski & Nilsson, 2001; Marra et al., 2019). Nonetheless, BMI-for-age remains a widely accepted measure in educational research and is supported by many researchers (Borga et al., 2018; Orsso et al., 2020; Özdemir, 2023). Hence, this measurement is acceptable and applicable, especially among students, where simplicity, cost-effectiveness, and ease of implementation are critical factors.

This study used the Students' Attitudes Toward Physical Education Scale (SPEA) to measure learning attitudes. One limitation of self-reported data, such as that collected through the SPEA, is the potential for social desirability bias, where students may provide answers that they believe are more socially acceptable than their actual opinions (Lavidas et al., 2022). However, this is a

common challenge in educational research. To minimise this bias, several strategies were implemented. First, the anonymity of the responses was emphasised to the students, which has been shown to reduce social desirability bias significantly (Lavidas et al., 2022).

Additionally, the survey was administered in a controlled, neutral environment, reducing the likelihood of external influences on the responses. These methods have been successfully applied in previous research and have proven effective in minimising bias (Cooper et al., 2020; Stantcheva, 2023). While self-reported data inherently carries some bias, these approaches ensure that the reliability and validity of the data collected remain high.

Finally, an important limitation of this study is the potential impact of growth spurts that may occur during the three-month intervention period. Junior high school students are at a critical stage of physical development, and natural changes related to growth, such as height, weight, muscle mass, and flexibility, may occur independently of the intervention. These physiological changes could affect outcomes such as body mass index, strength, agility, and flexibility, making it difficult to fully attribute the observed improvements within the group to the SEM program itself. However, the study used random grouping to minimize bias, and it also analyzed gender differences in the dependent variables and adjusted for family income as a covariate to further reduce bias.

1.8 Conceptual and Operational Definitions of Terms

1.8.1 Football

Conceptual Definition: Football, or soccer, is a competitive sport played between two teams of eleven players on a rectangular grass or artificial turf field. The objective is to score goals by working as a team to move the ball into the opponent's net. A standard match lasts 90 minutes, divided into two halves of 45 minutes each. The game is regulated by global rules established by FIFA, covering aspects such as the number of players, field dimensions, goal size, and fouls. As the world's most popular sport, football is prominently featured in global events like the FIFA World Cup and national leagues worldwide.

Operational Definition: In this study, 7-a-side football refers to a modified version of football, where each team consists of seven players. It is typically played on a smaller field than the regular 11-a-side version, with reduced pitch dimensions and smaller goalposts. The match duration is shorter, usually comprising two halves of 25 to 30 minutes each. Although the basic rules and technical skills are similar to 11-a-side football, the smaller playing area results in a faster-paced game, emphasising individual skills, short passing, and quick transitions between offence and defence. This study employs 7-a-side football as an intervention to assess its effects on students' physical fitness and learning attitudes.

1.8.2 Standard Teaching Approach

Conceptual Definition: The Standard Teaching (ST) approach, characterised by its teacher-centred methodology (Gubacs-Collins, 2015), is devised primarily by instructors with minimal or no student participation in the development of teaching plans and evaluation criteria (Mosston & Ashworth, 2008). During the teaching process, instructors maintain authority over the classroom by formulating regulations and behavioural norms that students must adhere to. Students' engagement and skill performance are closely monitored by instructors, who deliver corrective feedback with a critical eye on identified errors (Metzler, 2017). Consequently, students are tasked with repetitive exercises conducted within predetermined timeframes to achieve precise reproduction of the movement techniques taught by the instructor. This teaching model typically features a highly structured classroom setup (e.g., skills-drills, lines, or circuit organisation) and is deemed conducive to motor skill development during the early stages of learning, particularly for novice scholars, as it concentrates on honing motor skills through repetition and escalating difficulty (Xu et al., 2021).

Operational Definition: The Standard Teaching approach is often referred to as direct teaching, traditional teaching, or traditional style (Franco et al., 2021; Layne, 2015; Li et al., 2022; Pritchard et al., 2008; Wallhead & Ntoumanis, 2004). In China, the Standard Teaching approach, or the skill mastery teaching model, emphasises "the three basics" of teaching, encompassing basic knowledge, fundamental skills, and basic techniques. The teaching process is typically segmented into three sections, namely the beginning,

foundational, and concluding parts, leading some scholars to term it the "three-step" teaching model. This approach was developed and refined following the establishment of the People's Republic of China and drew from the teaching philosophy of the former Soviet Union. It has garnered favour among physical education teachers on the front lines (Kang & Gao, 2016; Gao, 2020; Zhuo, 2021).

1.8.3 Sport Education Model

Conceptual Definition: This approach is founded on SEM derived from game theory, serving as the guiding ideology. It places competition at the core of the teaching process, focusing on adapting competition rules and equipment. The instructional methods encompass direct guidance from teachers, cooperative learning, and partner learning. Moreover, it adopts fixed groups and role-playing as the organisational structure (Gao, 2020). The primary objective is to simulate vibrant and authentic sports scenarios to offer students genuine and immersive sports experiences across various skill levels. It maximises games' educational and instructional potential, striving to educate students within the most true-to-life situations, ultimately enabling them to become capable, knowledgeable, and passionate sports participants (Siedentop & Van der Mars, 2022).

Operational Definition: In junior high school (Level 4) football teaching, the operational definition of the SEM is specified by the following six characteristics:

a. Seasons: Football instruction is conducted according to seasons, typically encompassing an entire football season within a semester, from preliminary matches to finals. This ensures that students concentrate on football over an extended period, enhancing their interest and skills in the sport (Siedentop & Van der Mars, 2022).

b. Affiliation: At the beginning of the semester, students are assigned to fixed football teams, which remain constant. This stable team structure helps students develop a sense of team spirit and belonging (Siedentop & Van der Mars, 2022).

c. Formal Competition: Formal football matches are organised where students take on roles such as players, coaches, and referees. The matches are conducted according to standard football rules, helping students understand and practice the spirit of fair competition (Siedentop & Van der Mars, 2022).

d. Culminating Event: A culminating event, such as a championship or exhibition, is held at the end of the semester, marking the end of the football season. This event showcases the student's skills and progress and provides an opportunity for celebration and reflection (Siedentop & Van der Mars, 2022).

e. Record Keeping: Students' performances are documented, including technical statistics, match results, and individual progress. Each team's and individual records help students and teachers understand learning progress and areas for improvement (Siedentop & Van der Mars, 2022).

f. Festivity: At the end of the semester, a football festival is organised, including finals and award ceremonies, celebrating the students' efforts and achievements throughout the season. These activities enhance students' sense of participation and accomplishment (Siedentop & Van der Mars, 2022).

Therefore, this study will follow the six characteristics of the Sport Education Model to intervene in students' physical fitness and learning attitude.

1.8.4 Junior High School Students

Conceptual Definition: Junior high school students' study at the junior high school level, typically covering grades 6 to 9, with an age range of approximately 12 to 15. This aligns with the United Nations Educational, Scientific and Cultural Organisation (UNESCO) definition of the junior high school education stage. During this stage, students are in the early stages of adolescence, experiencing significant physiological and psychological changes. Junior high school education aims to provide students with foundational knowledge and skills while promoting social and emotional development (UNESCO, 2011).

Operational Definition: In the Chinese education system, "junior high school" is typically divided into the 7th, 8th, and 9th grades and lasts three years. Junior high school students" specifically denote individuals enrolled in junior high school who fall within the broader category of "teenagers." These students typically have an average age of 13 to 15.

1.8.5 Physical Fitness

Conceptual Definition: Physical fitness, as defined by Medrano-Ureña et al. (2020), refers to an individual's capacity to engage in daily activities with vitality and alertness, free from excessive fatigue. This capacity includes enjoying recreational pursuits and responding effectively to unforeseen

emergencies. It encompasses both aspects of physical fitness and skill-related attributes. Health-related components of fitness encompass body composition, cardiorespiratory fitness, flexibility, muscle strength, and muscular endurance. In contrast, skill-related physical fitness pertains to the motor capabilities necessary for speed, agility, power, balance, coordination, and reaction time, as Pedretti et al. (2014) described.

Operational Definition: At the junior high school level, the operational definition of physical fitness is assessed based on the standards outlined in the "Guidance Manual for the Physical Fitness Testing of Youth Football Players" issued by the Chinese Football Association and the Sports Science Institute of the General Administration of Sport in 2021. Physical fitness encompasses several key indicators: body composition (measured by BMI-for-age), flexibility (assessed through the sit and reach Test), agility (evaluated through the Arrowhead Run Test), speed (assessed by the 30-meter Sprint Test), upper body strength (evaluated through the Pull-up Test), lower body strength (assessed by the Vertical Jump Test), and cardiorespiratory endurance (evaluated through the YOYO and IR1 Test).

1.8.6 Physical Education Learning Attitude

Conceptual Definition: Physical education learning attitude is a distinct aspect of sports education. It represents the culmination of students' evaluations and internal emotional responses to physical education instruction. It reflects students' cognitive and affective perspectives on physical education. The affective component measures the degree of

emotional attraction or feeling toward an attitude object, and the cognitive component accounts for the beliefs about the characteristics of the attitude object (Gonzàlez, 1992). The attitude towards physical education learning is a pivotal determinant in guiding students' engagement in physical activities, serving as a significant psychological mechanism that directs their participation (Xie, 2006).

Operational Definition: This study employed Silverman's (2017) two-factor model of Attitude Theory, which encompasses cognitive and affective dimensions, to assess students' attitudes toward physical education. The measurement and evaluation of students' attitudes toward physical education were conducted using the Students' Attitudes toward Physical Education Scale (SPEA), developed for junior high school students by Subramaniam and Silverman (2000). The SPEA has been widely used and validated in multiple studies. For instance, Subramaniam and Silverman (2007) used the scale in an extensive survey of 995 middle school students to assess attitudes and found reliability and validity in its results. Furthermore, a 2013 study by Huanhuan Hu et al. examined the Chinese version of the SPEA, confirming its reliability (Cronbach's alpha = 0.90) and construct validity through confirmatory factor analysis. Both studies demonstrate the SPEA's robustness in measuring attitudes toward physical education across diverse student populations. Subsequently, scholars widely used it (Duan et al., 2015; Li et al., 2014). Hence, this is a suitable model for this study.

CHAPTER 2

LITERATURE REVIEW

This chapter provides an in-depth review and explanation of relevant literature in the following sections: 2.1 Football Players' Match Requirements and Characteristics; 2.2 Current Situation of Chinese Football Curriculum; 2.3 Definition of Terms; 2.4 Pedagogical Models in Physical Education; 2.5 Effects of Sport Education Model on Students' Physical Fitness; 2.6 Effects of the Sport Education Model on Students' Learning Attitude; 2.7 Other Factors Influencing Learning Attitude and Physical Fitness of Students; 2.8 Theoretical Framework; 2.9 Theoretical Framework Establishment and 2.10 Conceptual Framework.

2.1 Football Players' Match Requirements and Characteristics

As football continues to evolve, modern football players must meet various physiological, cognitive, and technical demands during matches. Whether in training or during competition, players' physical fitness is a primary focus to enhance their performance. To win a football match, it is essential to analyse the match demands and the characteristics required for football players to reach peak performance (Anderson & Sally, 2013; Xiao, 2022).

2.1.1 Football Match Requirements

Modern football matches require players to balance physical fitness and skills over extended periods of high-intensity activity. While the official match duration is typically 90 minutes, actual playtime may be extended due to

halftime breaks, injury stoppages, and added time. During these high-intensity intermittent activities, players typically cover distances of 10 to 12 kilometres, with frequent speed changes (Morgans et al., 2024; Dalen et al., 2021). High-intensity phases, such as sprints and sudden stops or changes in direction, are often interspersed with low-intensity activities like walking or jogging, making the physical demands of football highly complex (Caldbeck, 2020).

2.1.2 Characteristics of Football Players

The success of football players is closely linked not only to their psychological and technical abilities but also to their physical fitness. Studies show that speed and endurance play pivotal roles in football, with sprinting ability particularly crucial for offensive and defensive actions (Castillo et al., 2021). Balance and coordination are also fundamental skills, as the frequent changes in direction and positioning during matches make these abilities indispensable (Burhaein et al., 2020).

Modern football players' talent identification and development process often begins in youth, with national youth teams selecting and nurturing future talent through various training programs. These youth programs typically include comprehensive physical fitness assessments to evaluate young players' physical capacities, laying a solid foundation for their future entry into the national team (Shalar et al., 2021). Thus, schools play a vital role in developing students' football abilities as the foundation for selecting outstanding national team players.

2.2 Current Situation of Chinese Football Curriculum

China’s struggle in international football competitions, particularly in events like the FIFA World Cup, has sparked extensive national discourse, especially considering the country's achievements in other sports (Sun, 2022). Despite the government's significant efforts to promote football within the educational system and cultivate young talent, clear challenges have stunted the development of football at the grassroots level. Programs targeting elementary school students have successfully increased motivation and participation, but this involvement sharply declines upon students' transition to junior high school (Xue & Li, 2023).

One of the main issues lies in the standardised approach to physical education in Chinese schools, particularly in football curricula. The current football curriculum heavily emphasises structured knowledge transmission and technical skill acquisition, leaving limited room for creativity or decision-making by students (Ma, 2019). While beneficial for developing basic skills, this task-oriented focus often needs to encourage cognitive engagement and problem-solving, essential components of practical application in sport (Ma, 2019). The curriculum’s over-reliance on summative assessments, such as exams and performance evaluations, has been identified as contributing to negative student attitudes toward physical education (Gao, 2020). These rigid teaching methods reduce students' motivation for autonomous learning and limit opportunities to foster a lifelong engagement with physical activity, a critical concern during adolescence (Kang & Gao, 2016).

Furthermore, beyond the adverse effects on learning attitudes, there is growing concern over Chinese students' declining physical fitness levels. National fitness monitoring data highlights a continuous downward trend in student fitness over the past three decades (Dong et al., 2019), underscoring the urgent need for reforms in physical education curricula. To inspire greater involvement, a more flexible and student-centred approach is necessary to address students' physical and psychological needs.

2.3 Definition of Terms

This section introduces key concepts relevant to this study, including pedagogical models, physical fitness components, learning attitude dimensions, and characteristics of junior high school students.

2.3.1 Pedagogical Models

Pedagogical models offer diverse approaches to organising teaching and learning processes, particularly in physical education, where different models can help develop students' physical, cognitive, and social skills. Below are several widely used and recently researched models:

1) **Sport Education Model**

SEM provides students with experiences that mimic real sports participation, focusing on roles like players, referees, and coaches. This model fosters long-term participation in physical activities by enhancing student competence, enjoyment, and sports literacy (Siedentop et al., 2011; Bessa et al., 2021).

2) Teaching Games for Understanding

Teaching Games for Understanding (TGfU) emphasises learning tactical awareness and decision-making in game contexts rather than isolated technical skills. This model promotes greater game understanding and strategy development, improving engagement and learning outcomes (Pill & Gambles, 2023; Barba-Martín et al., 2020).

3) Teaching Personal and Social Responsibility

Teaching Personal and Social Responsibility (TPSR) aims to teach students life skills like respect, effort, and leadership through physical activities. It has been shown to impact student behaviour in class positively and their conduct in broader social settings (Havas, 2021; Pérez-Ordás et al., 2020).

4) Cooperative Learning

This model encourages students to work in groups to accomplish common learning goals, emphasising communication, collaboration, and peer support. It enhances students' social interaction and teamwork skills, crucial in sports and academic contexts (Bores-García et al., 2021).

5) Health-Based Physical Education

Health-Based Physical Education (HBPE) promotes lifelong physical activity and health rather than sports performance. It integrates health education with physical activities to ensure students understand the importance of fitness and well-being beyond the school environment (Bowler et al., 2022).

2.3.2 Physical Fitness

1) Body Composition

Body composition refers to the proportion of fat and lean mass in the

body. A commonly used measure for assessing body composition is Body Mass Index (BMI), which is calculated as weight (kg) divided by height squared (m^2). However, for individuals under 18, BMI-for-age is the recommended metric, as it accounts for age- and sex-specific variations in growth and development (WHO, 2023; CDC, 2022). In this study, BMI-for-age is used to assess the body composition of junior high school students, ensuring that growth-related variations are properly accounted for in the analysis. This approach provides a more accurate reflection of physical health in adolescents and aligns with internationally recognized assessment standards.

2) Flexibility

Flexibility involves the range of motion around joints, contributing to injury prevention and overall functional movement. The sit-and-reach test is a standard measure of this component (Baquet et al., 2001; Nokariya et al., 2023; Garramona et al., 2024).

3) Agility

Agility is the capacity to change direction quickly and accurately, an essential aspect of performance in many sports (Ardoy et al., 2011; Aphamis et al., 2015).

4) Speed

Speed is the ability to move swiftly over a short distance, typically evaluated through sprint tests (Stastny et al., 2023; Stankovic et al., 2023).

5) Upper Body Strength

This measures the ability of the upper muscles to exert force, often assessed through exercises like push-ups (Van Den Tillaar & Ball, 2020; Zalleg et al., 2020).

6) Lower Body Explosiveness

Lower body explosiveness refers to the power produced by the muscles in the legs, which is crucial for jumping and sprinting (Perez et al., 2022).

7) Cardiovascular Fitness

Cardiovascular fitness refers to the body's ability to deliver oxygen during prolonged physical activity, and it is commonly measured through endurance tests (de Almeida Andrade, 2020; Malshe & Gunjal, 2021).

2.3.3 Learning Attitude

Learning attitude refers to the mindset and disposition that students bring to the learning process, encompassing cognitive and affective dimensions.

1) Cognitive

This dimension refers to students' understanding, beliefs, and perceptions related to learning in physical education. It is critical in PE because students' comprehension of physical education concepts can significantly affect their overall learning outcomes (Hastie & Casey, 2014).

2) Affective

The affective aspect of learning attitude relates to students' emotions and motivation towards learning. A positive affective attitude can enhance engagement and long-term participation in physical activities (Shen & Shao, 2022).

2.3.4 Junior High School Students

Junior high school students, typically between 13 and 15, are at critical physical and psychological development stages. This period is characterised by rapid physical growth, cognitive advancements, and increased social awareness, making it an ideal time for interventions targeting physical fitness and learning attitudes (Casey et al., 2016; Zhu, 2022).

2.4 Pedagogical Models in Physical Education

In order to gain a comprehensive understanding of pedagogical models in physical education, the researchers conducted two review studies. One is an umbrella review, which focuses on individual pedagogical models, titled From Mind to Body: An Umbrella Review of Pedagogical Models from an Embodied Cognition Perspective, submitted to PLOS ONE, and its status is currently under review. The other is a review of hybrid pedagogical models integrated with the Sport Education Model, titled *Optimizing Learning Outcomes in Physical Education: A Comprehensive Systematic Review of Hybrid Pedagogical Models Integrated with the Sport Education Model.* This article was published in *PLOS ONE* (Zhang et al., 2024). Both studies conducted electronic searches of six databases in 2024: Web of Science, SCOPUS, EBSCOhost, PubMed, PsycINFO, and PsycARTICLES. The inclusion and exclusion criteria for the literature were applied to validate the results. The studies reviewed 25 systematic reviews on individual pedagogical models and 30 primary studies on hybrid pedagogical models. The detailed contents of both reviews are provided in Appendix A, Figure 1, Table 1 and Table 2,

and Figure 2, Table 4, Table 5, and Table 6. The results of the reviews identified the advantages of SEM for student learning outcomes and the gaps in the research.

Pedagogical models have garnered significant attention in physical education (PE) research due to their proven ability to enhance student learning outcomes across various domains, including substantial benefits to motor, cognitive, social, and affective domains. The following are two major research gaps identified in pedagogical models based on two reviews:

(a) Population Gap

It is worth noting that while a large body of research is concentrated in countries like Spain, the United States, and the United Kingdom, there is relatively less research on pedagogical models in physical education in some Asian countries and regions, especially in China.

(b) Intervention Gap

Whether used alone or in combination with other teaching models, SEM is the most researched topic, indicating its effectiveness as an intervention (Zhang et al., 2024; Fernandez-Rio & Iglesias, 2024).

2.5 Effects of the Sport Education Model on Students' Physical Fitness

To better understand the impact of SEM on students' physical fitness, the researcher conducted a systematic review, including health-related physical fitness and skill-related physical fitness. *The title is The Effect of The Sport Education Model on Sports Ability: A Critical Systematic Review*, published in Revista de Psicología del Deporte (Journal of Sport Psychology) (Zhang et al., 2023). A comprehensive search was conducted on websites such as

PubMed, Web of Science, EBSCOhost (CENTRAL, SPORTDiscus), Scopus, CNKI, and Google Scholar to select relevant literature. The literature chosen focused on objective measurements of physical fitness following the implementation of SEM. Based on the selection criteria, a total of twelve studies were selected, four of which were related to physical fitness. Overall, the research indicated that the SEM impacts students' physical fitness. PRISMA diagrams, detailed search strategies, and publication details are presented in the form of Figure 2, Table 4, and Table 5, respectively, in Appendix B. This section focuses on the results of SEM's impact on students' physical fitness in the context of teaching and highlights prominent gaps relevant to current research.

The systematic review concluded that SEM significantly impacts students' physical fitness development compared to Standard Teaching methods. However, by using the PICOS principle to analyse the previous literature, it can obtain the following gaps:

(a) Intervention Gap

The sources or basis of the arranged teaching content should be introduced when setting up the curriculum using SEM. It also needs to be determined whether the scheduled course content aligns with the participants' skill levels. Therefore, the current research lacks a systematic design for football courses in physical education colleges based on the "Chinese Compulsory Education Physical Education and Health Curriculum Standards (2022 Edition)."

Furthermore, a significant portion of the studies have class hours of less than 18 lessons (Hastie et al., 2009; Nikravan et al., 2019; Chen, 2017; Segovia & Gutierrez, 2020), which contradicts the

recommendations and requirements outlined in the "Chinese Compulsory Education Physical Education and Health Curriculum Standards (2022 Edition)", which advocate for setting up large-unit teaching with more than 18 lessons, which is beneficial for students' understanding of physical education knowledge systematically.

Additionally, some scholars propose that inappropriate intervention frequency and duration may be one of the reasons for the need for more significant improvement in students' physical fitness, leading to lower reliability and persuasiveness of the results (Yin, 2019). However, a review of the literature reveals that a significant portion of the studies have low intervention frequencies (Wu, 2008; Chen, 2017; Wahl-Alexander & Chomentowski, 2018; Nikravan et al., 2019) and short intervention periods, such as four weeks (Hastie & Trost, 2002; Pritchard et al., 2015; Ward et al., 2017) or five weeks (Hastie et al., 2009c).

Previous studies only described the positive impact of SEM intervention on students' physical fitness without incorporating relevant theories (Hastie et al., 2009; Nikravan et al., 2019; Wahl-Alexander et al., 2018; Zhang, 2020), which made it difficult to explain how SEM improved students' physical fitness fully.

(b) Outcome Gaps

There is a need to explore the effects of the SEM on a broader range of physical fitness indicators, many studies now focus ONLY on skill and limited physical fitness indicators, potentially overlooking other key dimensions of students' physical development which is important for thier growth.

2.6 Effects of the Sport Education Model on Students' Learning Attitude

To better understand the impact of SEM on student learning attitudes, the researcher conducted a systematic review, including cognitive attitudes and affective attitudes learning. The title is "*The Effect of the Sport Education Model in Physical Education on Student Learning Attitude: A Systematic Review*," which BMC Public Health published (Zhang et al., 2024). To select relevant literature, the researcher searched PubMed, Web of Science, EBSCOhost (CENTRAL, SPORTDiscus), Scopus, Google Scholar, and other websites. The chosen literature focused on objectively measuring student learning attitudes after implementing SEM. According to the selection criteria, a total of 13 articles were selected. Overall, the studies indicate that SEM can promote students' learning attitudes. Appendix C provides the PRISMA diagram, detailed search strategy, and publication details in Figure 3, Table 6, and Table 7, respectively. This section focuses on the results of SEM's impact on student learning attitudes in the teaching context and highlights prominent gaps in current research.

The systematic review presents compelling evidence supporting the positive impact of the SEM on students' attitudes toward physical education learning. By using the PICOS principle to analyse the previous literature, it can obtain the following gaps:

(a) Population Gap

Most of the literature included in this review predominantly focuses on high school and college students, with limited findings for other age groups. There is an urgent need for comprehensive research

initiatives that prioritise investigating the impact of SEM on attitudes toward physical education learning across diverse age groups, especially junior high school students. This approach aims to foster a more inclusive understanding of SEM's effectiveness. It ensures its benefits are explored and validated across various educational stages, laying a solid foundation for evidence-based instructional practices in physical education.

(b) Intervention Gap

The study highlights the need for more thorough examinations of model fidelity in specific investigations, raising concerns about the reliability and validity of observed improvements reported in both SEM and ST groups before and after the experiment. Future research should prioritise rigorous fidelity checks to enhance the credibility and interpretability of research findings when comparing the effectiveness of different instructional models. Due to the complexity of students' learning attitudes, there are differing views on the definition of learning attitude.

However, previous studies have merely described the positive impact of SEM intervention on students' learning attitudes without incorporating relevant theories (Viciana et al., 2020; Luna et al., 2020; Chu et al., 2022; Iserbyt et al., 2023) and lack explanations of the dimensional components of students' learning attitudes. This makes explaining how SEM comprehensively improves students' learning attitudes difficult. Moreover, most studies during the intervention process did not provide comprehensive and detailed information on the frequency, intensity, time, and type of the intervention (Wallhead & Ntoumanis, 2004; Spittle & Byrne, 2009; Perlman, 2010; Menickelli & Hastie, 2014; Wallhead et al., 2014; Cuevas et al., 2016; Xu et al., 2019), leading to doubts about the scientific validity and rationality of previous intervention plans.

(c) Comparison Gap

The systematic review revealed that 27 studies used a single experimental group design without a control group (Zhang et al., 2024). The included studies also often relied on quasi-experimental or qualitative methods, needing more evidence from randomised controlled trials (RCTs) (Zhang et al., 2024).

(d) Outcome Gap

Through the literature review, it is recognised that there are problems in using tools for evaluating learning attitudes in previous studies, and there needs to be scales or questionnaires developed based on attitude theory to investigate students' learning attitudes in SEM teaching intervention. Previous studies lacked scales based on Attitude Theory to assess students' learning attitudes. Instead, they used intrinsic motivation, intention to be physically active, enjoyment, and fun to reflect learning attitudes. According to proponents of the two-component attitude theory, Attitude Theory defines attitude as an individual's emotional and cognitive (positive or negative) evaluation of an attitude object. Therefore, using these indicators to reflect students' learning attitudes may not accurately measure their attitudes. Thus, using the SPEA scale, based on Attitude Theory as the theoretical framework, is recommended to measure students' attitudes. This approach would ensure that the evaluation of learning attitudes is more comprehensive and theoretically grounded, providing a clearer understanding of how SEM impacts students' attitudes toward learning football.

2.7 Other Factors Influencing Learning Attitude and Physical Fitness of Students

Many studies have investigated the factors that influence students' physical fitness. The most critical factors that significantly affect students' physical

fitness are family income (Suyati et al., 2022), time spent in physical activities during and outside of school hours (Hajian-Tilaki et al., 2011), and growth and nutritional status (Gokhale & Kirschner, 2003; Kail & Cavanaugh, 2007; Elamin, 2005; Jeejeebhoy, 2000; Narayan & Khan, 2007).

2.7.1 Family Income

Family income is a critical factor influencing students' participation in physical education and sports activities. Research suggests that children from higher-income families tend to have greater access to sports facilities, professional coaching, and extracurricular physical activities, which can contribute to better physical fitness and more positive attitudes toward sports (Huang et al., 2023; Wang et al., 2022). Conversely, students from lower-income backgrounds may face financial barriers that limit their opportunities for sports participation, leading to lower engagement in physical education (Chen et al., 2021).

In the context of SEM implementation, family income can influence students' ability to engage effectively in SEM-based physical education. Unlike traditional teacher-centered approaches, SEM emphasizes student autonomy, teamwork, and competition, often requiring students to take on roles such as team captains, referees, or coaches (Bessa et al., 2021). However, students from lower-income backgrounds may have fewer prior experiences with organized sports or access to structured extracurricular activities, which can impact their initial adaptation to the SEM framework (Zhang et al., 2024).

Recent studies have shown that students from lower-income families may also have limited access to sports-related knowledge, equipment, and support from their families, affecting their confidence and motivation to participate in student-led activities (Li et al., 2023; Sun & Liu, 2022). This suggests that when implementing SEM, additional support mechanisms, such as structured guidance from teachers or peer mentoring programs, may be necessary to ensure equitable participation across different socioeconomic backgrounds. The family income of the students in this study can be queried in the student file data.

2.7.2 Time Spent in Physical Activities during and Outside of School Hours

Studies have shown that regular physical activity positively affects students' physical development, physical fitness, and motor skills development (Active Living Research, 2007; Lauer, 2008). However, failure to achieve 60 or more minutes of moderate-to-vigorous intensity physical activity per day is a significant proportion of school children (Trost & Loprinzi, 2008). It is a global problem (Guthold, 2019). Studies have shown that students in China also suffer from the same problem. From 2001 to 2016, the proportion of boys aged 11-17 who lack physical activity was above 80%, while the proportion of girls who lack physical activity was higher than that of boys (Guthold et al., 2020). Wee (2008) conducted a study on the impact of physical activity on students' physical health among middle school students in a school in Malaysia. It improved students' participation in physical activity by enhancing their awareness of physical education courses. After the four-week

intervention, physical fitness tests (push-ups, sit-ups, sit-forward bends, four-quadrant jumps, and 1500m endurance runs) were conducted. By comparing the pre-test and post-test data, it is concluded that active participation in sports can significantly improve students' physical fitness.

2.7.3 Growth and Nutritional Status

Body growth is usually the increase in body tissue mass from infancy to adulthood, and the leading indicators include height, weight, body fat percentage, and muscle mass. Some scholars have proposed that the most significant bone and muscle fibre changes occur between 12 and 17 (Kail & Cavanaugh, 2007). Therefore, BMI-for-age is often used to measure physical growth. Yang et al. (2022) studied the relationship between physical fitness index (PFI) and BMI-for-age of middle school students and concluded that BMI-for-age value was negatively correlated with exercise score, and then proposed to reduce the incidence of overweight and obesity and improve physical fitness by enhancing students' exercise enthusiasm.

Zhang and Xie (2022) analysed and compared the physical quality of young football players with different BMI-for-age levels and concluded that abnormal BMI-for-age values of male athletes would hurt the power, flexibility, speed, and dexterity of lower limbs, while low BMI-for-age values would hurt the upper limb pulling muscle group of male athletes. High BMI-for-age harms the abdominal muscle endurance of female athletes. There was a positive correlation between BMI-for-age and upper limb muscle strength and lower limb explosive strength in male athletes and a negative correlation between

BMI-for-age and aerobic endurance, speed, and dexterity. A negative correlation exists between BMI-for-age and upper limb strength and abdominal strength endurance quality in female athletes. In addition, a study of trends in physical fitness indicators by growth and nutritional status from 1985 to 2014 revealed That overall physical fitness was 167% lower between 1995 and 2014. Also, except for sit-and-reach, the other fitness components significantly declined from 1995 to 2014 ($p<0.01$), particularly forced vital capacity and endurance running. In addition, there were significant differences in physical fitness among students with different nutritional status. (Dong et al., 2019).

2.8 Theoretical Framework

2.8.1 Introduction of Sport Education Model

The SEM comprises six core characteristics: seasons, affiliation, formal competition, culminating event, record keeping, and festivity (Siedentop et al., 2020) (See Figure 2.1). Collectively, these elements foster students' physical fitness and improve their attitudes toward learning football.

(a) Season

A sports season requires an extended period that allows students to develop skills, understand concepts, and build excitement through progressively challenging activities. The sports season typically consists of preseason, formal competition, and playoff. This structure encourages a gradual upward progression, with fewer drills and practices as the season continues and more emphasis on competitive games (Siedentop et al., 2020), enabling students to gradually increase physical activity intensity, which contributes to cardiovascular fitness and muscular strength (Gómez-Mármol et al.,

2021). Over time, students build self-confidence and resilience, fostering a more positive learning attitude as they witness physical and skill development (Zhang et al., 2024; Opstoel et al., 2020; Bessa et al., 2019).

(b) Affiliation

Affiliation refers to forming fixed groups where students work together as a team. Team cooperation is a vital feature of the Sport Education Model, and every student is expected to develop a sense of teamwork (Siedentop et al., 2020). During the teaching process, physical education teachers can group students in various ways, whether randomly or based on abilities. Group members collaborate to practice techniques and strategies to help their team succeed (Spittle & Byrne, 2009). Working as a team promotes accountability and cooperation, which are essential for maintaining a positive attitude towards learning football. Furthermore, as students engage in collective physical activities, their physical fitness levels improve through increased participation in varied group-based drills and exercises (Zhang et al., 2023). Roles such as scorers, referees, or team managers allow students to understand the sport more, further developing cognitive learning outcomes and their attitude toward teamwork and collaboration (Sun, 2018).

(c) Formal Competition

Formal competition motivates students to engage fully in both practice and gameplay. This competitive environment encourages students to apply their learned skills and strategies in real-time situations, which helps refine their physical abilities, particularly agility, coordination, and stamina (Hastie et al., 2011; Spittle & Byrne, 2009). Competitions also provide opportunities for self-assessment, which can improve self-discipline and perseverance. Students can measure their progress and effort relative to their peers, positively impacting learning attitudes (Sun, 2013).

(d) Culminating Event

The culminating event, often in the form of a final match, reinforces student participation by creating a celebratory and meaningful conclusion to the season. This final event maximises student engagement, fostering a sense of accomplishment that enhances motivation and enjoyment (Gao, 2020). From a physical perspective, preparing for a culminating event encourages sustained physical effort and participation throughout the season, leading to improved fitness (Zhuo, 2021). By focusing on a tangible goal, students remain motivated to participate, which positively influences their learning attitudes toward the sport.

(e) Keeping Records

Record-keeping provides students with clear evidence of their progress, contributing to cognitive and physical development. Regular performance tracking enables students to set goals and monitor their improvements, which has been shown to enhance intrinsic motivation (Spittle & Byrne, 2009). This element also encourages students to reflect on their performance, promoting a sense of accountability that improves their physical fitness and learning attitudes (Sun, 2013). For example, students who see measurable improvement in their physical abilities often develop a more positive attitude toward ongoing participation in football (Cohen et al., 2020).

(f) Festivity

Festivity celebrates student achievements and the social aspects of learning through sports. Acknowledging students' efforts and improvements in a public setting, such as award ceremonies or celebrations, the SEM curriculum promotes positive reinforcement, closely linked to improved learning attitudes (Sun, 2018). Celebratory events create a sense of belonging and accomplishment, motivating students to continue physical training and contribute to team success. This sense of festivity fosters enjoyment and excitement

around the sport, which ultimately enhances both the physical and affective learning outcomes (Spittle & Byrne, 2009).

In conclusion, the six core characteristics of the SEM work synergistically to create a positive learning environment, making it a potentially effective method for improving students' physical fitness and learning attitudes.

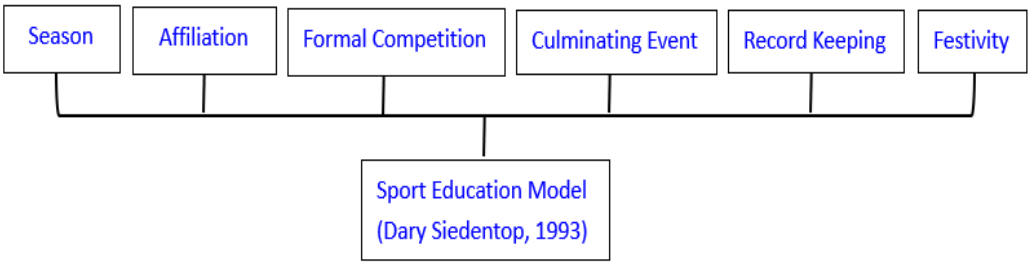


Figure 2.1 : The Characteristics of the Sport Education Model
(Source: Dary Siedentop, 1993)

2.8.2 Embodied Cognitive Theory

Unifying investigators of embodied cognition is the idea that the body or the body's interactions with the environment constitute or contribute to cognition in ways that require a new framework for its investigation, emphasising the importance of the subject's body in cognitive ability (Shapiro, 2019), its role is to explain the causal relationship between brain, body, and environment, namely: The framework of embodied cognition allows for a more direct causal link between the actor's life experience and the cooperative dynamics of the brain, body, and environment (Wilson & Golonka, 2013) (See Figure 2.2). Cognition is physically interactive, embedded in physical contexts, and manifested in physical bodies.

From the perspective that phenomenologists favour, however, all cognition is embodied, interactive, and embedded in dynamically changing environments (Shapiro, 2019). The concept of an organism understanding its environment depends on the nature of its body, so different embodied organisms appreciate their environment in various ways. When the agent integrates into an appropriately designed physical or social environment, the cognitive "load" required for the task can be reduced. For example, Martin and Schwartz (2005) found that when children were allowed to manipulate a pie chart, they were more successful at calculating a quarter of an eight than when they only watched the pie chart. The cognitive load required to sail a large naval ship is beyond the capabilities of any one person but can be divided among many specialists, each with their own specific task (Hutchins 1996). It is easier to search for a particular title by alphabetising the books on the shelf than by simply placing them randomly on the shelf. In all these cases, individuals' cognitive abilities are enhanced when they have the opportunity to interact with features of an appropriately organised physical or social environment.

Embodied Cognitive Theory provides a robust framework for explaining how SEM may promote students' physical fitness and learning attitudes. SEM combines students' physical actions, social interactions, and environmental factors in a real-world, contextualised learning environment, aligning closely with the emphasis of embodied cognition on the interaction between physical abilities and the environment (Shapiro, 2019). In SEM programs, students enhance their physical fitness through actual football skills training and

establish cognitive engagement and positive learning attitudes through physical interactions in team collaboration and competitive settings.

Embodied cognition theory suggests that students, through embodied activities and interaction with natural environments, increase their motivation to learn, thereby improving learning outcomes (Wilson & Golonka, 2013). For instance, students participating in simulated games and practices in authentic sports environments can enhance their physical fitness and foster a more positive attitude toward learning football by integrating body and environment. Thus, the real-world environment provided by SEM helps students translate theoretical learning into practical skills, enhancing their interest and confidence in sports (Shapiro, 2019). These factors collectively contribute to improvements in physical fitness and learning attitudes, providing a theoretical explanation of how SEM may promote students' physical fitness and learning attitudes from the perspective of Embodied Cognitive Theory.

In SEM, students engage in authentic sports scenarios where they take on roles such as players, referees, or coaches, requiring them to make real-time decisions based on their physical engagement. These experiences align with Embodied Cognition Theory's principle of sensorimotor integration, where learning is enhanced through active, embodied participation rather than passive knowledge acquisition (Beilock, 2015).

For example, in a football-based SEM curriculum, students do not merely memorize tactical strategies but physically enact them in competitive game

settings, reinforcing decision-making and problem-solving through real-time embodied experiences. This dynamic interaction between movement, cognition, and environment fosters deeper learning, making SEM an effective approach to enhancing both physical fitness and cognitive engagement in sports education (Rudd et al., 2021).

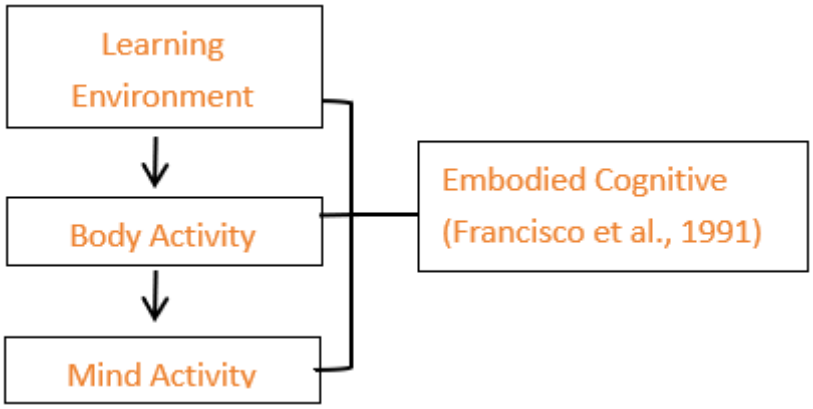


Figure 2.2 : Embodied Cognitive Theory (Source: Francisco et al., 1991)

2.8.3 Attitude Theory

In the past few decades, there have been various definitions of attitudes. Some scholars believe that attitudes are structures people form based on existing information (Schwarz & Bohner, 2001), while others propose that attitudes can be considered a stable internal tendency (Allport, 1935). Attitudes have also been conceptualised in terms of the number of constituent components. The first perspective is the single-component view, which regards attitude as a singular structural concept, typically focusing on the emotional response to the object of the attitude (Fazio & Zanna, 1981; Insko & Schopler, 1972; Ajzen & Fishbein, 2005). However, some scholars argue

that using only "emotion" to define attitude limits its scope (Subramaniam & Silverman, 2000).

The second perspective is the two-component view, which emphasises the cognitive and emotional evaluation of the object of attitude (Bagozzi & Burnkrant, 1979; Gonzàlez, 1992; Mohsin, 1990; Oppenheim, 1992). Bagozzi & Burnkrant (1979) compared the validity of a one-component attitude model with a two-component conceptualisation, concluding that attitude is a complex construct that should include both cognitive and affective dimensions (See Figure 2.3).

The third perspective is the three-component view, which includes not only cognitive and emotional evaluation but also motivation and behavioural tendencies toward the attitude object (Hilgard, 1980; Reddy & LaBarbera, 1985; Triandis, 1971; Krech et al., 1962). However, some researchers (McGuire, 1989; Seeman, 1993) questioned this three-dimensional notion, finding that the attitude measure explained only about 10% of the variation in behaviour. Studies showing higher correlations focus on attitudes and behavioural intentions rather than actual behaviour (McGuire, 1989). This paper adopts the two-component view proposed by Rosenberg in 1960, which posits that attitudes consist of an affective component (positive or negative feelings about the problem) and a cognitive component (beliefs or thoughts about the issue) (Rosenberg, 1960).

Attitude Theory suggests that students' engagement in physical education is shaped by their cognitive beliefs and emotional responses (Eagly & Chaiken, 1993). This theory is particularly relevant in SEM-based interventions, where student attitudes toward participation, teamwork, and self-improvement influence their overall learning experiences.

In SEM, students develop cognitive learning attitudes by engaging in strategic decision-making, rule comprehension, and self-assessment. For instance, a student serving as a team captain must analyze game tactics, organize team strategies, and make leadership decisions, reinforcing higher-order cognitive skills within the physical education setting (Perlman, 2012).

The affective learning attitude in SEM is strengthened through positive emotional experiences, social interaction, and a sense of accomplishment. Unlike traditional physical education models that may induce anxiety due to performance-based grading, SEM fosters a supportive team environment where students experience enjoyment, motivation, and confidence (Fernández-Río & Cecchini, 2016). For example, when students celebrate a successful team strategy in a football match, their emotional attachment to the learning process is reinforced, leading to greater motivation for continued participation.

By integrating Embodied Cognition Theory and Attitude Theory, this study highlights how SEM-based physical education enhances both cognitive

engagement and positive learning attitudes, offering a comprehensive approach to student development in junior high school settings.

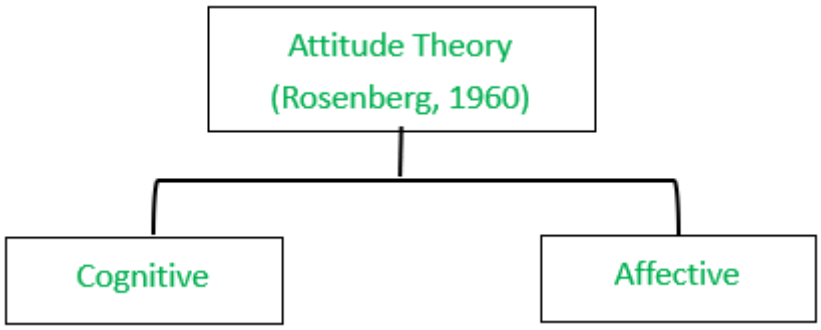


Figure 2.3 : Attitude Theory (Source: Rosenberg, 1960)

2.9 Theoretical Framework Establishment

Numerous researchers have affirmed that the gaming-centred teaching model provides students a more engaging learning environment (Siedentop et al., 2011; Harvey & Jarrett, 2014). Within this framework, the SEM, characterised by its six features (season, affiliation, formal competition, culminating event, record keeping, festivity), offers students rich learning experiences (Siedentop, 2011; Bessa et al., 2019).

Embodied cognition emphasises the significant causal relationship between the environment, body, and mind (Wilson & Golonka, 2013). By altering the external learning environment to impact bodily activities, learners' internal cognitive processes are influenced (Chiel & Beer, 1997).

Attitude theory can be regarded as an individual's stable internal psychological predisposition towards objects (Allport, 1935). Its

dual-component view emphasises that attitudes comprise cognitive and affective evaluations (Bagozzi & Burnkrant, 1979; Mohsin, 1990; Zajonc & Markus, 1982).

Therefore, the theoretical framework of this study is established as follows: SEM offers enriched learning contexts and environmental experiences, facilitating bodily activities in students and consequently influencing their cognitive and affective processes. Through SEM, Embodied Cognitive and Attitude Theory elements are integrated to formulate a teaching program, and a balanced teaching intervention plan is established based on the FITT (frequency, intensity, time, and type) principle (Demers, 2015). This intervention impacts students' learning performance (learning attitude and physical fitness) (Chen & Gu, 2018). The learning attitude indicators are divided into cognitive and affective categories based on attitude theory. Additionally, research has shown that changes in bodily activities affect physical fitness (Ward et al., 2017), which is also a significant educational objective in teaching implementation (Carriedo et al., 2020). Hence, this study considers it a vital measurement indicator (See Figure 2.4 - Theoretical Framework).

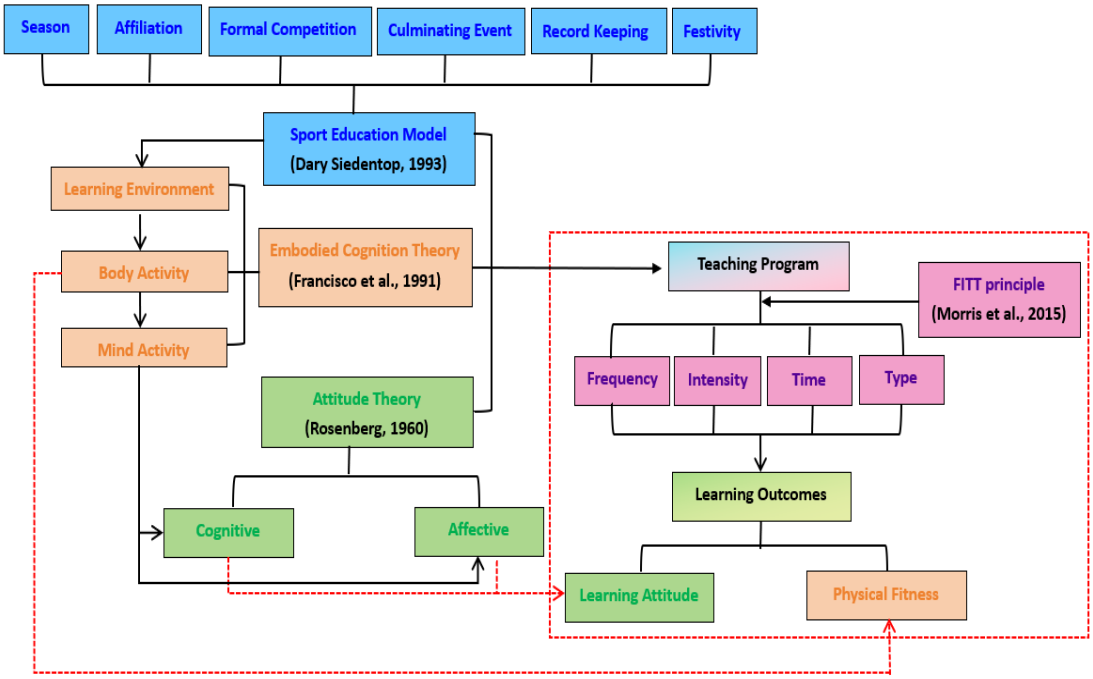


Figure 2.4 : Theoretical Framework

2.10 Conceptual Framework

In this study, SEM is the independent variable. The dependent variables are learning attitude (cognitive and affective) and physical fitness (cardiovascular fitness, flexibility, agility, upper body strength, lower body explosiveness, speed). Figure 2.5 illustrates the conceptual framework of this experimental study. It briefly overviews the SEM intervention treatment, while the control group received ST without additional intervention. Both groups underwent 36 sessions lasting 40 minutes each, three times a week, for 12 weeks.

2.10.1 Introduction to the FITT Principle

The FITT principle is a guideline to help individuals structure their physical activity programs for optimal health and fitness outcomes. FITT stands for Frequency, Intensity, Time, and Type, the four key variables that can be

adjusted to create an effective exercise routine. This principle is widely applied in general fitness and sports training programs, offering a flexible yet systematic approach to enhancing physical performance and health-related fitness.

Frequency refers to how often physical activity is performed. In physical fitness, this usually means the number of weekly exercise sessions. Depending on individual fitness goals and levels, frequency can vary. For instance, cardiovascular training might be recommended 3 to 5 times a week, while strength training might be advised 2 to 3 times a week (Jakicic, 2019).

Intensity describes how hard a person works during physical activity. Intensity can be measured in various ways, such as heart rate, perceived exertion, or specific performance targets (e.g., lifting a certain weight or running at a particular pace). It is crucial to balance intensity to ensure that exercise is challenging but not overwhelming, as excessive intensity can lead to injury or burnout (Mann et al., 2013).

Time refers to the duration of each exercise session. For general health, it is often recommended that individuals engage in at least 30 minutes of moderate-intensity exercise most days of the week. However, the time may need to be adjusted for more specific goals. For example, endurance athletes might need to train for more extended periods. In contrast, high-intensity interval training (HIIT) sessions may be shorter in duration but higher in effort (Karlsen et al., 2017).

The type of exercise indicates the kind performed, such as aerobic (running, swimming), anaerobic (weightlifting), or flexibility training (yoga). The type of exercise should align with the individual's goals. For example, a person aiming to improve cardiovascular health would benefit from aerobic exercises, while someone focusing on muscle strength should incorporate resistance training (Wibisana & Royana, 2023).

The FITT principle allows individualised exercise programs adaptable to changing fitness levels and goals by manipulating these four variables. It is a fundamental concept in exercise science and is often applied in physical education, personal training, and sports performance enhancement.

2.10.2 Physical Fitness

Physical fitness is a multifaceted ability measured by various factors, often assessed through five health-related and six skill-related indicators in sports. Good physical fitness generally reflects an individual's ability to work effectively, enjoy leisure time, maintain health, resist hypo-motor diseases, and respond well to emergencies. Health-related physical fitness includes body composition, cardiovascular fitness, flexibility, muscular endurance, and strength (Corbin & Lindsey, 1997). Body composition refers to the relative percentage of muscle, fat, bone, and other tissues in the body, with a moderate body fat percentage being optimal for health. Cardiovascular fitness pertains to the efficiency of the heart, blood vessels, and respiratory system in delivering nutrients and oxygen to muscles during sustained activity. Flexibility refers to the range of motion in joints, muscular endurance

relates to a muscle's ability to perform repetitive actions without fatigue, and strength reflects the muscle's capacity to exert force or lift heavy objects.

Skill-related physical fitness, closely associated with athletic performance, comprises agility, balance, coordination, power, reaction time, and speed. Agility is the ability to change body movement direction quickly and accurately, which is advantageous in skiing and gymnastics. Balance refers to maintaining stability while still or moving, which is essential for activities such as unicycling or balance beam performances. Coordination involves using sensory input from different body parts to execute motor tasks smoothly and accurately, such as in fancy rope jumping or basketball. Power is the ability to convert energy into force quickly, as seen in throwing or weightlifting. Reaction time is the interval between a stimulus and the initiation of a response, which is crucial in sports like sprinting. At the same time, speed reflects the capacity to act rapidly over a short period, as required in races.

After reviewing gaps in previous research using the PICO principle, this study selected the following physical fitness indicators: cardiovascular fitness, flexibility, coordination, and balance. These indicators are relevant to the study's goal of assessing the impact of the SEM on students' physical fitness outcomes.

2.10.3 PE Learning Attitude

Proponents of the two-factor model argue that attitudes include cognitive and emotional aspects and have long been considered valid and reliable

(Silverman, 2017). The affective component measures the degree of emotional attraction or feeling toward an attitude object, and the cognitive component accounts for the beliefs about the characteristics of the attitude object (Gonzàlez, 1992). Some scholars set “usefulness” as the main component of the cognitive dimension in the score verification of students’ assessment tools for physical education attitude, such as usefulness teacher (UT), usefulness curriculum (UC), and usefulness peer (UP). Set “enjoyment” as the main component of the emotional dimension, such as enjoyment teacher (ET), enjoyment curriculum (EC), and enjoyment peer (EP). Reliability and validity analyses supported the two-component view of attitudes. The scores of the enjoyment and usefulness factors showed good psychometric properties in reliability and validity (Subramaniam & Silverman, 2000). Some scholars also refer to “satisfaction” and “comfort” as emotional dimensions and believe that “satisfaction” is a general emotion, and “comfort” is a specific emotion, such as relaxation or anxiety related to PE. The students’ perception of PE teachers was set as a cognitive dimension. And the scale has excellent psychometric properties (Orlić et al., 2017). Student Attitudes Towards Physical Education Scale -The SPEA scale was developed by Subramaniam and Silverman in 2000 and is considered one of the most thorough, empirically based, and contemporary valid instruments (Subramaniam & Silverman, 2000).

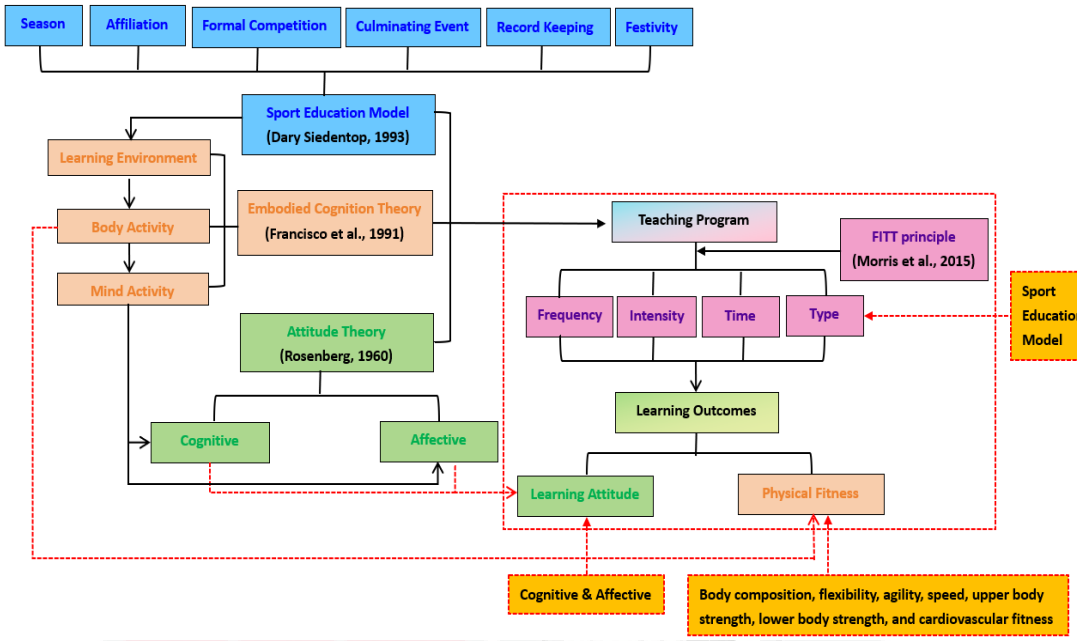


Figure 2.5 : Conceptual Framework

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter describes the detailed curriculum implementation steps and research design adopted to examine the effects of SEM in physical education on Chinese junior high school students' physical fitness and learning attitudes. In addition, the related issues involved in data collection and coarse processing are discussed: the sampling process, study site, equipment and instrumentation, reliability, data collection procedures, and the methods and purposes of data analysis.

3.2 Research Design

This study is a pre-test and post-test true experimental design consisting of a control group and an intervention group: the control group (ST) and the experimental group (SEM). This study's independent variable (IV) is the football's SEM curriculum, which was implemented as an intervention in the experimental group. The dependent variables (DV) are the physical fitness components, including body composition, flexibility, agility, speed, upper body strength, lower body strength, cardiorespiratory endurance, and the students' learning attitudes towards football. These variables were measured to determine the impact of the SEM intervention on physical performance and cognitive/affective dimensions of learning attitudes. The specific experimental process is shown in the figure below:

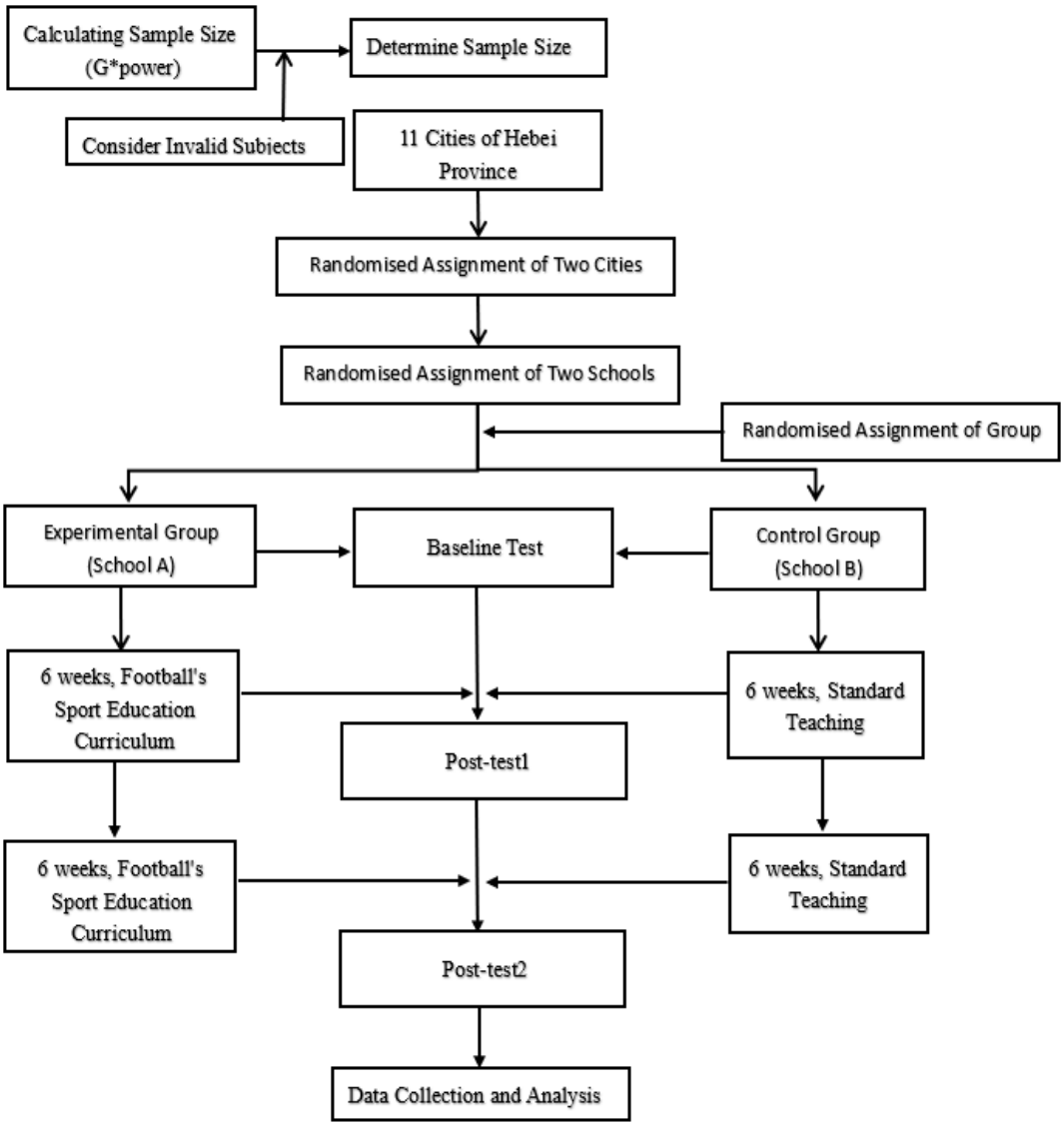


Figure 3.1 : Research Flow

3.3 Study Location

Population Distribution: Hebei Province is one of the most populous provinces in China, with a diverse mix of urban and rural areas. This diversity in population distribution allows for a broad selection of research subjects, representing different regions and demographics, thereby enhancing the study's representativeness.

Geographic Location: Hebei Province is located in northern China, with distinct climate and geographical characteristics compared to other regions. These geographical differences may influence students' physical fitness levels and learning attitudes. Hence, choosing Hebei Province as the research location helps to ensure the study results' generalisability.

Educational Resources: Hebei Province boasts abundant educational resources, including school facilities and teaching staff. Disparities in these resources may affect the quality of physical education curriculum and learning outcomes. Therefore, conducting the study in Hebei Province allows for a better understanding of the impact of physical education curriculum interventions on students.

Cultural Background: Hebei Province has a rich history, and cultural heritage could influence students' participation in sports and attitudes towards learning. By studying students in Hebei Province, we can gain insights into how cultural factors impact the effectiveness of physical education curriculum interventions.

In summary, the selection of Hebei Province as the research location is based on considerations of population distribution, geographic location, educational resources, and cultural background, ensuring the representativeness and generalizability of the study results.

3.4 Population and Sampling Method

Clarifying the target population, sample size calculation, selection criteria, and sampling method are essential to ensuring this study's experimental effect.

3.4.1 Target Population

Identify the experimental educational setting: Educational experimentation is typically conducted within the framework of school curricula, considering the pragmatic aspects of experiment execution and the desire for results that closely align with authentic pedagogical contexts. The study focuses on students aged approximately 13 to 15 attending junior high schools in various urban and rural areas of Hebei Province.

3.4.2 Sampling Size Calculation

In order to reflect the gender diversity inherent in mixed-class instructional settings within Chinese schools, thoughtful consideration of male and female student participation is imperative. The sample size was determined using the G* power 3.1.9.7 analytical method outlined by Faul et al. (2009) to ensure robust statistical power (See Figure 3.2).

Before calculating the sample size, the effect size of our intervention was based on the SE teaching intervention study conducted in previous studies (Wahl-Alexander & Chomentowski, 2018; Nikravan et al., 2019; Ma, 2012;

Pritchard et al., 2015) calculated (Table 3.1). Based on these results, the minimum effect size of Power (f=0.19) was calculated (Ma, 2012).

Table 3.1 : Effect Size Calculation for Research Variables (Mean and SD)

Variables	Baseline	Post-test	Effect Size (f)	Reference
Cardiovascular fitness	53.83 (19.58)	86.96 (24.64)	0.60	Wahl-Alexander & Chomentowski (2018)
Flexibility	38.17 (4.10)	42.17 (3.40)	0.47	Nikravan et al. (2019)
Agility	10.87 (0.91)	10.26 (0.63)	0.36	Nikravan et al. (2019)
Upper body strength	8.39 (4.04)	9.88 (3.81)	0.19	Ma (2012)
Lower body explosiveness	32.9 (11.4)	37.9 (10.8)	0.22	Nikravan et al. (2019)
Speed	7.52 (0.35)	7.12 (0.34)	0.50	Nikravan et al. (2019)
Cognitive	15.03 (4.67)	19.34 (4.95)	0.40	Pritchard et al. (2015)
Affective	22.26 (5.00)	26.48 (2.93)	0.46	Luna et al. (2020)

Since the study consisted of six sets of repeated measurements (twice), the F-test - ANOVA was used. Considering type error I (α)=0.05, power (1-β) =0.80, two measurements, and two groups, the sample size was calculated using repeated measurements with interplay. The following is a sample size calculation based on G*Power, with each group having a sample size of n= (48/2) = 24.

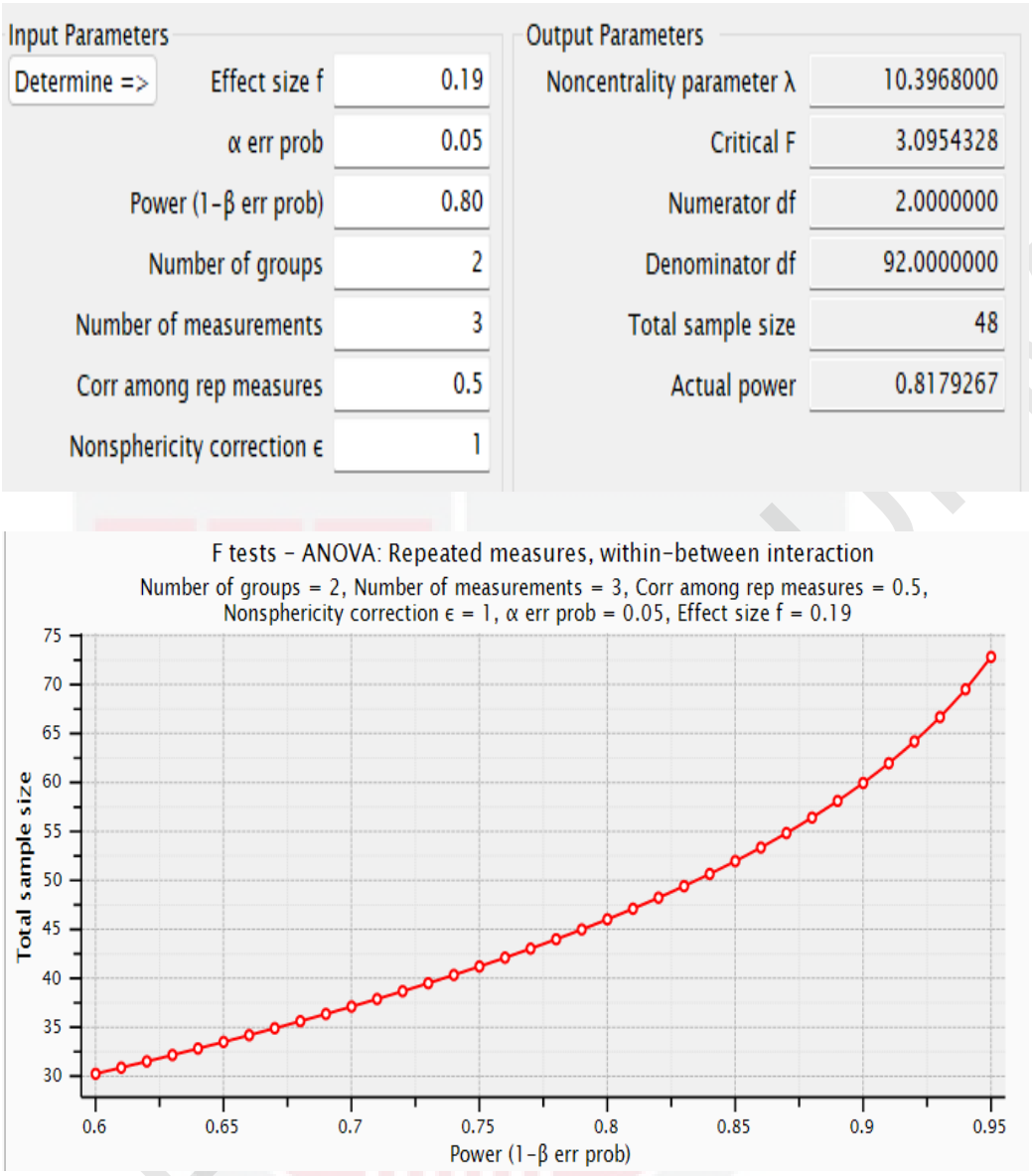


Figure 3.2 : Sample Size Calculation Based on the Repeated Measure within-between Interaction

This study is a cluster-randomised controlled trial. Donner et al. (1981) suggested that to achieve the desired statistical performance under Cluster Randomized Controlled Trial (CRCT), the individual randomisation assumed by the designed sample size may be amplified by the Design Effect (DE) (Donner et al., 1981). A priori, a sample size requirement of n=48 for the total sample size and n=48/2=24 for each group was determined (G*Power

3.1.9.7). p (ICC) values in human studies range from 0.01 to 0.02 (Murray et al., 1994; Murray & Short, 1995; Murray and Short, 1996).

The following equations were provided to calculate the sample size:

Equation 1:

$$DE = 1 + p\,(m - 1)$$

Equation 2:

$$ESS = DE * mk$$

Where m=number of subjects in a cluster, k-number of clusters, mk=total number of participants in a clustered study, ESS=effective sample size, DE=design effect, and p=intra cluster correlation coefficient (ICC) (Killip et al., 2004b). Thus, in this study, the total adequate sample size:

$$ESS = [1 + 0.01 * (24 - 1)] * 24 * 2 \approx 60$$

Considering the 20% probability of invalid subjects or dropouts (Smedegaard, 2016; Cramer, 2015), the number must be at least 72. Since the study was

divided into two groups, the sample size was 72 participants, with n =36 per group.

3.4.3 Criteria for Sampling

Inclusion criteria: Participants must be enrolled in junior high school, including male and female students. They must be in good health and capable of engaging in regular physical education activities. All participants and their parents have provided written informed consent for their involvement in the study.

Exclusion criteria: Participants failing to complete all tests due to attendance-related issues (25%) (Albaladejo-Saura et al., 2020) or illness during experimental teaching or testing will be excluded.

3.4.4 Sampling Method

In this study, 98 students volunteered and were screened from two schools, 49 in each, covering grades 7-9, with very similar numbers of students in each grade. Three participants needed to meet the inclusion criteria (screening failure). Five participants were excluded due to schedule conflicts or guardian objections, and two were excluded due to non-payment of campus accident insurance (see Figure 3.3). At this time, the number of students enrolled in each school was 44 and assigned to the SEM group (n=44) or ST group (n=44), and written informed consent was obtained from the participants (signed by both the participants and their parents or guardians). Two students in the SEM and ST groups were lost to follow-up

during the experiment. The main reasons included that the students could not complete the course for their reasons, and they needed rest to transfer to other schools and get pneumonia. The total dropout rate was 4.55%. In the end, the SEM and ST groups had 42 participants each, no students had an absenteeism rate of more than 25%, and all students completed all the test content and were ready for the final analysis.

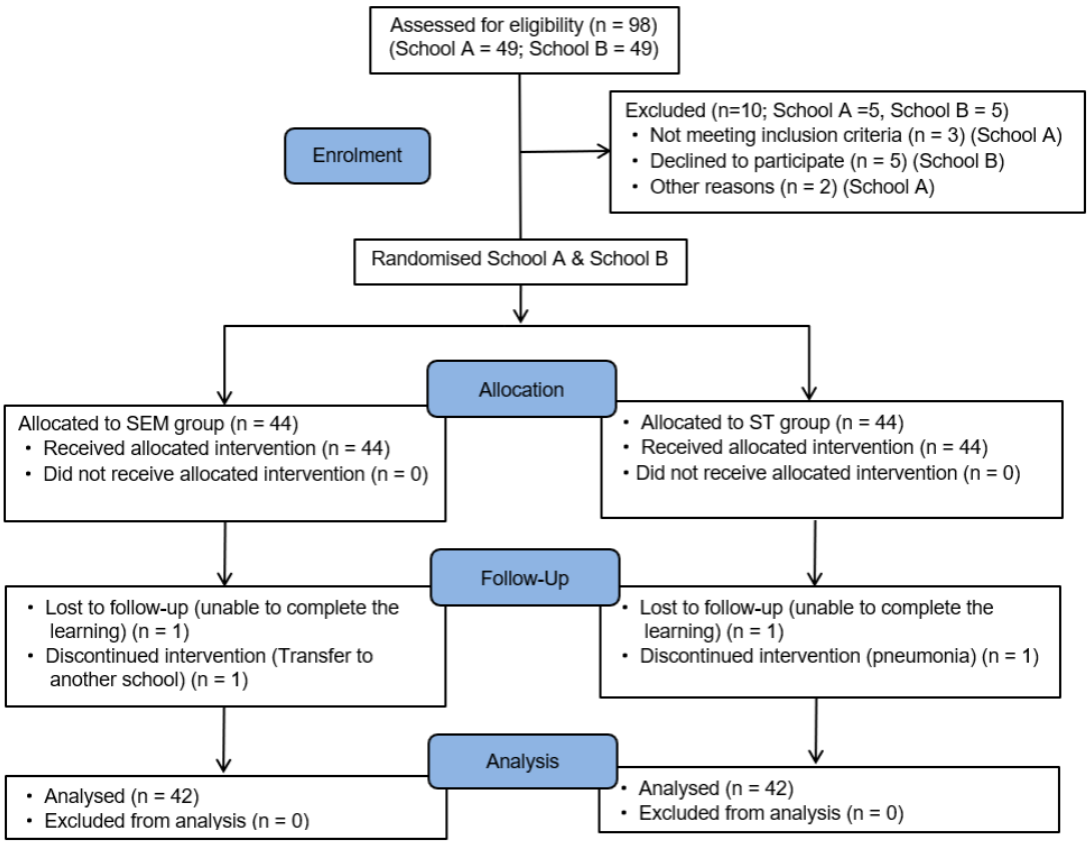


Figure 3.3 : Participants' Flow Diagram
(Source: Moher et al., 2012)

3.5 Instruments

The present study employed three distinct instruments for the measurement of control variables, covariates, and dependent variables: (i) Socio-demographic Information, (ii) Physical Activity Questionnaire for

Adolescents (PAQ-A), and (iii) Students' Attitudes Toward Physical Education Scale (SPEA).

3.5.1 Socio-demographic Information

The tool for gathering socio-demographic information is derived from Anderson's Sport Activity Behaviour Model, which principally encompasses several key elements: individual characteristics, including socio-economic background, cultural attributes, gender, age, health status, among others, which influence individuals' interests and levels of engagement in sports activities; decisions regarding sports activity participation, where individuals, based on their own characteristics and environmental conditions, decide whether to engage in sports activities, with these decisions potentially influenced by factors such as time, money, and resources; ecological factors, including the social and cultural environment, family environment, community facilities, schools, and workplaces, which affect the accessibility and feasibility of sports activities for individuals; and individuals' sports activity participation behaviours, which entail the actual engagement in sports activities, encompassing frequency, type, duration, etc. (Alkhawaldeh et al., 2023). Pilot studies assessed reliability, while validity was evaluated based on expert recommendations before application.

3.5.2 Physical Activity Questionnaire for Adolescents (PAQ-A)

The current study utilised the Physical Activity Questionnaire for Adolescents (PAO-A), developed by Li Xin et al., to assess the physical activity levels of school students. This instrument is suitable for self-administered and

retrospective survey questionnaires for children and adolescents aged 7-18. The questionnaire consists of nine items. The first item (Q1) primarily evaluates adolescents' moderate to vigorous physical activity (MVPA) over the past week. Item two inquires about adolescents' activity levels during physical education classes; items 3-6, respectively, ask about the number or frequency of days or times students engage in active physical exercise during lunch break, after school in the afternoon, evenings, and weekends. Item seven requires adolescents to provide a rough summary of their extracurricular activities over the past seven days, while item 8 entails self-assessment of daily physical activity over the past week.

Additionally, the questionnaire includes item 9, which asks respondents whether any particular circumstances affected their regular participation in physical activities over the past seven days. Each item comprises five options, with scores ranging from 1 to 5 based on the overall activity level from low to high. The total score is the average score of the first eight items, with higher scores indicating higher physical activity levels. Classification criteria based on PAO scores are outlined in Table 3.2. Details for the Chinese versions are shown in Appendix D.

Table 3.2 : Physical Level Division

Levels	Score
Low-intensity activity level	<2
Moderate-intensity activity level	2-3
High-intensity activity level	>3

3.5.3 Students' Attitudes toward Physical Education Scale (SPEA)

SPEA (English version) has a total of 20 items, and its measurement purpose is to evaluate students' affective and cognitive of physical education (Subramaniam & Silverman, 2000). Affective and cognitive were labelled enjoyment and usefulness, respectively. Both enjoyment and usefulness were assessed for the two sub-domains of PE teachers and curriculum. A 5-point Likert scale ranging from "strongly disagree" to "strongly agree" (1 - 5) was used, with a possible total score of 100 for overall attitudes. The positive attitude increased with the increase in the score. Scholars verified the tool to be consistent with the validity and reliability of psychometric evidence (Subramaniam & Silverman, 2000). The scale has been translated into multiple versions, including Chinese (Hu et al., 2014). This study used the Chinese version because the participants were Chinese students. Details for the English and Chinese versions are shown in Appendix E (see Appendix E for Permission).

3.5.4 Assessment and Method for Physical Fitness

Promoting students' physical development has gradually become the primary goal of any physical education curriculum (Kamil et al., 2012). However, it can be challenging to assess physical fitness with a single test item (Miller, 2006). The current study usually assessed health status through a series of tests. Chinese Football Association and the China Institute of Sport Science jointly issued the Youth Football Ability Test Guidance Manual in March 2021 as this study's physical fitness assessment method. This assessment measures six

aspects of physical fitness: flexibility (Sit- and- Reach test), agility (Arrowhead Agility test), speed (30-meter Sprint test), muscle endurance (Sit-up test), lower body explosive power (Standing Long Jump test), and cardiovascular fitness (YOYO test, using the IR1 standard for ages 13-15). The list of specific test items is shown in Table 3.3. The test methods and instruments of the project are shown in Appendix F. These instruments have good reliability and validity, as indicated by the literature.

Table 3.3 : Physical Fitness Assessment Items and Methods

Variable	Sub-type	Testing methods	CI 95%/ Reliability
Physical fitness	Body composition	BMI-for-age (Özdemir, 2023)	0.995
	Flexibility	Sit and reach (Aboshkair, 2012)	0.970
	Agility	Arrowhead Agility test (Chalil et al., 2018)	0.995
	Speed	30-meter Sprint (Wisløff et al., 2004)	0.710
	Upper body strength	Pull-up (Aboshkair, 2012)	0.980
	Lower body explosiveness	Vertical Jump (Keiner, 2022)	0.940
	Cardiovascular fitness	Yo-Yo test (IR1) (Bangsbo et al., 2008)	0.998

3.6 Validation of Questionnaires

The three questionnaires included PAQ-A, Socio-demographic information, and SPEA.

3.6.1 Validity of the Questionnaire

3.6.1.1 Validity of PAQ-A

The Chinese version of the PAQ-A utilised by the researchers was translated and improved by Doctor Li Xin's team from the Department of Sports Rehabilitation at Beijing Sport University based on the English version of the PAQ-A questionnaire. Internal consistency reliability of the questionnaire was assessed using Cronbach's alpha coefficient, while test-retest reliability and criterion validity were examined using Spearman rank correlation. The study demonstrated acceptable reliability and validity. Within a specific range, the Chinese version of the PAQ-A questionnaire is suitable for monitoring physical activity among Chinese adolescent samples and has been widely used by researchers in China (Li et al., 2015; Cao et al., 2018; Cai, 2023; Su et al., 2023).

3.6.1.2 Validity of Socio-demographic Information

Six professionals familiar with the terminology used (i.e., experts in sports) were invited to serve as panel members (details are provided in Table 3.4). The Item-Content Validity Index (I-CVI) analysis indicated that all items achieved acceptable levels of content validity within the range of relevance features (1-CVI=0.833-1.000, Kappa=0.816-1.000). Clarity of all items also reached acceptable levels (1-CVI=0.833-1.000, Kappa=0.816-1.000). Therefore, these variables demonstrate high content validity (For details about the questionnaire content and results, see Appendix G).

Table 3.4 : Personal Details of Questionnaire Experts

Organization	Title & Degree	Area for Research
• Yuncheng University	Associate Professor	Exercise Physiology
• Shangqiu Normal College	Associate Professor	Physical Education
• Yuncheng University	Associate Professor	Physical Education
• Jiangsu Normal University	Doctor of Philosophy	Exercise Physiology
• Yancheng Normal University	Associate Professor	Physical Education
• Yuncheng University	Associate Professor	Physical Education

3.6.1.3 Validity of SPEA

The Chinese version of the Students' Attitudes Toward Physical Education (SPEA) Scale was employed in this study, which Ho and his team translated (reference). Hu and his team were from the Laboratory of Sports Epidemiology at the Graduate School of Sport Sciences, Waseda University. The translation was based on the original scale developed by Dr. Subramaniam and Dr. Silverman in 2000 (Subramaniam & Silverman, 2000). The scale has been found to possess good reliability and validity (Cronbach's alpha = 0.90, ICC = 0.94). Additionally, it targets junior high school students, similar to the participants in this study (Hu et al., 2014) (see Appendix G for details).

3.6.2 Content of Validity of Physical Fitness Test

Content validity, about the robustness of score interpretation of a tool, refers to the extent to which these scores measure the content they purport to measure. Statistical analysis of the content validity of the developed instrument was conducted using the Item Content Validity Index (I-CVI).

Kappa correction coefficients were determined to assess the consistency between I-CVI and its correlation (Polit et al., 2006). Developers of scales typically consider 0.78 as the acceptable lower limit for I-CVI (Vadlin et al., 2015). A minimum value of 0.40 for Kappa is deemed sufficient for consistency (McHugh, 2012). Six professionals familiar with the terminology (i.e., experts in sports) (see Table 3.5) were invited to serve as panel members. Analysis of I-CVI indicated that the content validity of all items fell within an acceptable range, demonstrating relevance features (I-CVI=0.833-1.000 and Kappa=0.816-1.000). Therefore, these variables exhibit high content validity. (see Appendix G for details)

Table 3.5 : Personal Details of Questionnaire Experts

Organization	Title & Degree	Area for Research
• Hunan Normal University	Professor	Exercise Training and Conditions
• Huzhou College	Associate Professor	Strength and Conditioning Training
• Huzhou College	Associate Professor	Strength and Conditioning Training
• Yancheng Institute of Technology	Associate Professor	Exercise Physiology
• Zhejiang Yuexiu University	Associate Professor	Strength and Conditioning Training
• Huzhou Normal College	Associate Professor	Strength and Conditioning Training

3.7 Intervention

The intervention plan is developed based on guidelines from the American College of Sports Medicine (ACSM), the 2022 version of the National Standards for Physical Education and Health Curriculum, Level Four Football

Teaching Content, and previous research. Additionally, the intervention plan was reviewed and validated by six experts. The first step involved applying the FITT principles recommended by ACSM for adolescents (Frequency, Intensity, Time, and Type). The second step included training program types, comprising SEM and ST. The final component focused on intervention efficacy.

3.7.1 Developing the Intervention Program

The American College of Sports Medicine (ACSM) recommends that adolescents adhere to the FITT principles during exercise and provide progressive principles for adolescents.

"F" stands for weekly training frequency, set to three times per week in this study, which is a standard exercise frequency but has been rarely applied in SEM teaching research (Ludin, 2013; Wallhead et al., 2014; Trist et al., 2014; Guijarro et al., 2020). Therefore, a frequency of three times per week was adopted.

"I" represents training intensity. In physical education for adolescents, activities should start from low to moderate intensity and gradually increase to high intensity. Generally, students are advised to engage in activities within 50% to 85% of their maximum heart rate to achieve optimal health and fitness benefits (Stratton, 1996). Hence, this study utilised an intensity of 50% to 85% of maximum heart rate. As the teaching progresses, the intensity of exercise continues to rise.

"T" stands for "duration of intervention". Each session in this study lasted for 40 minutes. This duration was determined based on the fixed class length set by the government and schools at the experimental site. Therefore, the intervention duration for both the experimental and control groups was the same. A total of 12 weeks is a standard total duration in teaching and training interventions (Bai et al., 2023; Li et al., 2023; Petrušič et al., 2022), and combined with the intervention frequency of three times per week, totalling 36 intervention sessions, this meets the recommended intervention length by Sidentop et al. (Sidentop et al.,1991). Thus, this study adopted a 12-week intervention.

"T" represents the type of exercise used in this study: SEM.

3.7.2 Intervention Teaching Program

The intervention lasted 12 weeks, consisting of 36 football training sessions (three sessions per week, each lasting 40 minutes). This duration was chosen based on previous research demonstrating that 6 to 12 weeks is sufficient to observe significant improvements in physical fitness and learning attitudes among adolescents participating in structured physical activity programs (Lubans et al., 2012; Hastie & Casey, 2014). Several studies have indicated that interventions shorter than 6 weeks may not provide sufficient time for physiological adaptations or attitudinal shifts to occur (Fairclough & Stratton, 2006). On the other hand, extending the intervention beyond 12 weeks may yield further improvements, but it also increases the risk of student dropout,

motivational decline, and logistical challenges in school-based settings (Hollis et al., 2017).

In the context of Sport Education Model-based programs, previous meta-analyses have shown that a minimum of 6 weeks is typically required to influence students' engagement, motivation, and physical performance, while optimal effects are generally observed between 8 to 12 weeks (Bessa et al., 2019; Méndez-Giménez et al., 2022). Therefore, the 12-week intervention in this study aligns with best practices in physical education research and provides sufficient time for measurable changes in both fitness and learning attitudes.

The intervention lasted 12 weeks, with sessions conducted thrice weekly, consisting of 40 minutes of moderate to vigorous-intensity teaching activities (at 50-85% of HRmax). The participants were divided into an SEM group and a control group. Selection criteria for teaching teachers: Teachers must willingly and actively participate in the study and possess prior experience 10 years as junior high school football instructors. Both experimental groups received training and guidance from the teaching staff before the intervention, emphasising the requirements of course characteristics.

The content of each group's practice mainly included warm-up, primary content, and cool-down. The teaching content for the experimental and control groups was based on the recommended Level Four teaching content in the National Standards for Physical Education and Health Curriculum

(2022 edition) at the compulsory education stage. The comparison of the intervention teaching programs between SEM and standard teaching is summarised in Table 3.6. The teaching for both programs ensures a structured approach to warm-ups, skill acquisition, and cool-downs, with the main distinction being the focus on team-based learning and competition in the SEM versus more teacher-directed, individual practice in the standard teaching approach. The detailed Intervention Teaching Program follows a cluster randomised controlled trial (CRCT) registered at ClinicalTrials.gov (Approval No.: NCT06322459). The details are shown in Appendix H - Protocol.

Table 3.6 : Summarised Intervention Teaching Program

Component	Sport Education Model (SEM)	Standard Teaching
Warm-up (10 min)	Light jogging, dynamic stretching, agility drills with football ball, bodyweight exercises, ball control drills.	Same as SEM: Light jogging, dynamic stretching, bodyweight exercises, and ball control drills.
Basic Content (20 min)	Team-based learning and practice of football techniques such as passing, dribbling, stealing, and defense. Weekly focus on specific skills.	Teacher-directed learning, focusing on individual skill practice with the same football techniques as SEM. Weekly skill focus.
Week 10	Formal competition with 5v5 team matches, score recording.	Teacher-directed scrimmage with group divisions for 5v5 matches.
Week 11	Culminating event with 5v5 matches score recording.	Teacher-directed scrimmage similar to Week 10.
Week 12	Festivity, team skill demonstration, awards ceremony, and captain/teacher summaries.	Review and course summary with personal skill displays teacher summary.
Rest and Hydration (5 min)	Structured rest, hydration breaks, and light stretching.	Same as SEM: Structured rest, hydration, and stretching.
Cool-down (5 min)	Light jog, static stretching, deep breathing, reflection, and hydration.	Same as SEM: Light jog, static stretching, deep breathing, reflection, and hydration.

3.7.3 Validation for Intervention Teaching Program

The phrase "content validity" refers to determining the extent to which the dimensions and aspects of a concept can be defined correctly (Matore & Khairani, 2015). Specifically, experts clarified, comprehensive, and professionalised the equipment under study; that is, experts provided commentary on the issues that were eventually included in the equipment and instruments.

Six-panel experts, including three physical educators, two sports trainers, and one physical education psychologist, validated the intervention to guarantee that the content generated is suitable, effective, and safe for teenagers (the specific content of the questionnaire is shown in Appendix H - Protocol). The experts proposed two suggestions: In terms of the relevance and consistency of the intervention, both groups should have the same class time and must include the model fidelity procedures. The specifics are shown in Table 3.7.

Table 3.7 : Personal Details of Intervention Teaching Program Experts

Organisation	Title & Degree	Area for Research
• Nanjing Sport Institute	Professor	Physical Education
• Shangqiu Normal University	Associate Professor	Physical Education
• Huzhou College	Associate Professor	Sports Training
• Huzhou College	Associate Professor	Sports Training
• Haigang District 19th Middle School	Associate Professor	Physical Education
• Shihezi University	Associate Professor	Psychology of Physical Education

This study focuses on the intervention content. The results of the CVI analysis indicate that all items exhibit acceptable content validity index ranges for relevance features (CVI=0.833-1.000, Kappa=0.816-1.000). Clarity for all items also reached acceptable levels (CVI=0.833-1.000 and Kappa=0.816-1.000). Therefore, the intervention plan demonstrates high content validity. Specific details are provided in Appendix I (Table 1).

3.7.4 Control of Extraneous Variables

3.7.4.1 Ways to Control Variables before Intervening

To ensure the validity of the study's findings, several extraneous variables were controlled throughout the research process. These included individual differences in age, gender, baseline physical fitness levels, and prior football experience, which could potentially influence the outcomes of the intervention.

A cluster-randomized controlled trial (CRCT) design was employed, ensuring that students from different grades were randomly assigned to either the SEM or Standard Teaching (ST) group. By using this method, the distribution of extraneous variables such as age and gender was balanced across groups, minimizing systematic bias. The specific screening process is that we randomly select a class in each grade and randomly draw an equal number of boys and girls (8 boys and 8 girls) from each class. This method was used in both experimental group and control group.

In addition, before the intervention began, we conducted baseline tests on control variables (socio-demographic information, physical activity) and dependent variables (physical fitness, learning attitudes) to assess whether the indicators of both groups were consistent before the experiment.

Finally, we also investigated whether gender affected the dependent variable and further analysed the gender difference between groups in the analysis. At the same time, family income was used as a covariate to analyse some indicators.

3.7.4.2 Ways to Control Variables during Intervening

Based on the results of a random sample of experimental sites, all training and testing locations are arranged in Qinhuangdao City, Hebei Province. The training locations are as follows: the SEM group is located at the 19th Middle School, while the ST group is located at the 15th Middle School. The considerable distance between these two schools effectively prevents potential contamination between samples.

Intervention Timing: The intervention for both groups was conducted during the third period, which lasted 12 weeks and included Monday, Wednesday, and Friday mornings.

Training Content: All participants undergo the same warm-up, primary content, and cool-down procedures. Specific details are provided in Appendix H, Protocol.

Teaching Staff: One teacher and two assistants work in the experimental group, while one teacher and two work in the control group.

Qualifications of Teaching Staff: All teaching staff are physical education teachers at the respective schools, possessing extensive teaching experience with 12 years of teaching tenure. They hold intermediate-level professional titles and have master's degrees. All teaching staff are female teachers to minimise any adverse effects on the experiment due to variations in teaching levels.

Teaching Experimental Equipment: The district education department uniformly provides the teaching equipment at both schools, fully meeting the teaching requirements.

Teaching Control: Model Fidelity Test. Two investigators, familiar with each teaching approach, randomly observed lessons to assess the fidelity of the teacher's instructional model. The observers utilised a checklist created by Hastie et al. (2013), which prompted the observer to determine whether the instruction followed the instructional features of SEM or ST (see teaching checklist in Appendix J).

Other Controls: During the experimental period, participants in both the experimental and control groups were discouraged from engaging in any other form of exercise, seeking professional assistance, or participating in any activities to improve physical fitness to minimise the potential adverse effects of different activities on the experiment.

Environmental Control: Environmental factors can significantly impact on-site testing, especially outdoors, primarily due to weather conditions. Consideration should be given to the potential effects of weather changes on the testing environment (such as wet/slippery/muddy surfaces) and the participants (such as extreme temperatures). Tests should be conducted under very similar conditions to facilitate comparison. Therefore, this study considers environmental conditions when recording test scores and may adjust the testing time in extreme conditions.

3.7.4.3 Control for Test Protocol

Before testing, participants are briefed on the procedure and guidelines to ensure active participation with the correct attitude and cooperation with the testers throughout the process. Additionally, testers should insist that participants adhere to warm-up exercises before testing and relaxation practices after testing.

The testing protocol is a crucial factor influencing the accuracy of test data. Tests should be conducted strictly according to standardised protocols during the testing process. Otherwise, meaningful data may not be generated. Therefore, in the data collection process of this study, the same tester will conduct tests on participants according to the testing protocol. The testers will be blinded, meaning they will not know the names and groups of the participants. Additionally, testers will be trained and tested according to the Youth Football Player Fitness Test Guidelines issued by the Chinese Football

Association and the Sports Science Research Institute of the General Administration of Sport of China (see Appendix H, Protocol).

3.7.4.4 Control for Co-variate

Although genetics can influence health and physical fitness, recent research still cannot demonstrate these factors, mentioning that genetics is complex and not easily explainable (Bouchard et al., 2015). To identify and control covariates to ensure the effectiveness of the experimental design, a teaching experiment study indicated that Time spent in physical activities during and outside of school hours, family income, and nutritional status influence adolescents' physical fitness (Aboshkair, 2012). Therefore, these covariates will be statistically analysed in this study, and differences between the two groups were examined before the experiment.

3.8 Pilot Study

A pilot experiment was conducted to determine the reliability and effectiveness of the tool and the feasibility of the intervention procedure, as well as to familiarise the researchers with the test protocol and instrument. Before the pilot experiment, we obtained ethical approval (See Appendix H-Protocol). Through the analysis and discussion of this information, a solution idea can be provided for the reasonable solution to the problems arising in the experiment, which lays the foundation for the final version of the experimental method. Baseline and post-tests were conducted in the experimental group (no control group) to investigate the influence of football physical education curriculum intervention on junior middle school students' physical fitness and

learning attitude. The study included 42 randomly selected students from a seventh-grade science class at a school that was not included. Data were collected using inclusion and exclusion criteria for formal experiments. The intervention lasted for two weeks, with three lessons per week. The starts before the first week, and the post-test data is collected after the last class of the second week.

3.8.1 Reliability for Questionnaires

Internal consistency (internal reliability) and test-retest reliability (external reliability) are two critical measures to test the reliability of questionnaires (Hu et al., 2014). Internal consistency tests the homogeneity of each item in its dimension in the subscale of the questionnaire. Test-retest reliability refers to the stability of the results of repeated measurements of an instrument over some time. Only a questionnaire that satisfies internal and external reliability is valid (Hu et al., 2014). Internal consistency is usually measured using Cronbach's α coefficient for the internal consistency of the subscale of the questionnaire, with a value greater than 0.6 considered acceptable. Test-retest reliability was determined using the Pearson Correlation coefficient. More than 0.6 is generally recommended. For the results of the reliability of the questionnaires, please See Table 3.8-3.10.

Table 3.8 : Test-retest Reliability of the Socio-demographic Information (n=42)

Questionnaire	List of Questions	Intra-class Correlation Coefficient Cronbach's Alpha	95% Confidence Interval	
			Lower Bound	Upper Bound
Socio-demographic Information	Height	1.00	1.00	1.00
	Weight	1.00	1.00	1.00
	Family income level	0.989	0.980	0.994
	Family Cultural background	0.952	0.909	0.974
	Health concept 1	1.000	1.000	1.000
	Health concept 2	1.000	1.000	1.000
	Health behaviour	0.924	0.860	0.960

Table 3.9 : Cronbach's Alpha Value of the Students' Learning Attitudes toward Physical Education Scale

Questionnaire	Dimensions	Cronbach's Alpha	List of Questions	Cronbach's Alpha if item Deleted
Learning Attitudes	Affective	0.811	Q1	0.774
			Q2	0.760
			Q3	0.805
			Q5	0.780
			Q9	0.790
			Q11	0.798
			Q12	0.848
			Q15	0.817
			Q19	0.767
			Q20	0.789
	Cognitive	0.826	Q4	0.826
			Q6	0.813
			Q7	0.801
			Q8	0.845
			Q10	0.800
			Q13	0.830
			Q14	0.797
			Q16	0.784
			Q17	0.812
			Q18	0.778
	Total	0.804	NA	NA

Table 3.10 : Test-retest Reliability of the Students' Attitudes toward Physical Education Scale

Questionnaire	Dimensions	Intra-class Correlation Coefficient Cronbach's Alpha	95% Confidence Interval	
			Lower Bound	Upper Bound
Students' Attitudes toward Physical Education Scale	Affective	0.821	0.691	0.900
	Cognitive	0.819	0.687	0.898
	Total	0.809	0.672	0.893

3.8.2 Reliability of Physical Fitness Test

Reliability usually refers to the degree to which a test or instrument gives consistent results over repeated tests, i.e., consistency of measurement results. For the reliability measurement of the measuring instrument base in the pilot study, the reliability of the instrument was evaluated by the retest reliability ICC (0.758-1.000). All (n=42) of the instrument results are acceptable, and all are greater than 0.7; see Table 3.11 for details.

Table 3.11 : Measurement Reliability of the Physical Fitness Tests

Physical fitness	Indicators	Intra-class Correlation Coefficient	95% Confidence Interval	
			Lower Bound	Upper Bound
Body composition	BMI-for-age	1.000	0.999	1.000
Flexibility	Sit and reach	0.979	0.962	0.989
Agility	Arrowhead agility test	0.739	0.563	0.851
Speed	30-meter sprint	0.758	0.592	0.862
Upper body strength	Pull-up	0.955	0.991	0.997
Lower body explosiveness	Vertical jump	0.919	0.854	0.956
Cardiovascular fitness	Yo-Yo test (IR1)	0771	0.612	0.870

3.9 Research Permissions

Before the experimental school was selected, the agreement was signed after communicating with the superintendents of the two schools and obtaining their consent. The protocol allows researchers to collect and use data in research and protect student privacy. In addition, the parents were informed that the school would be teaching the SEM. The teaching procedures for this activity were explained, and their consent was obtained. The parents were later given the intent Consent Form (ICF) to sign if they agreed to allow their child to participate in this study. The ICF form was returned to research for safekeeping.

3.10 Data Collection

The pilot study will commence in November 2023, comprising three weekly football sessions. The experimental group will follow the SEM plan, while the control group will adhere to the Standard Teaching method. Both groups will receive identical courses in number and content, aligned with the football teaching content outlined in the "Compulsory Education Stage Physical Education and Health Curriculum Standards (2022 edition)". Each class will be scheduled during the third or fourth period in the morning to minimise any potential influence on teaching effectiveness and test results due to varying time intervals.

Before the intervention, besides organising participants to sign informed consent forms and testing the homogeneity of demographic variables, the experimental and control groups will undergo baseline testing in the first week

(Huang, 2000). This baseline testing will include assessments using PEAS and physical fitness tests. The test data will provide preliminary diagnoses of the subjects' physical abilities and learning attitudes, serving as the baseline scores before the experiment. By comparing the baseline data with the post-test data after the experiment, we can determine whether both groups of students have experienced changes in their levels before and after the experiment, thereby inferring the effect of the experiment.

Following 33 lessons, the experimental and control groups will undergo post-testing, with questionnaires and physical fitness assessments administered in the same class. If some participants fail to complete all tests due to excessive workload or illness, their data should be excluded during experimental teaching or testing.

3.11 Data Analysis

The experimental data were analysed statistically using Statistical Software for Social Science (SPSS) version 23. The significance level for the entire process was set at $p < 0.05$. This study's essential data analysis methods were conclusive analysis and descriptive analysis. Before data analysis, data cleaning and hypothesis testing were conducted. The researchers performed data entry and utilised descriptive techniques for quality control, including identifying data outliers and coding errors.

Descriptive statistics were utilised for all variables. Descriptive statistics were presented as mean and standard deviation (SD) or median and interquartile

range (IQR) for continuous or categorical variables (based on their distribution) and frequencies and percentages for categorical variables. Each variable was described, including characteristics, demographic data, and study variables (Grove, 2018).

The normality of data distribution was assessed to determine the type of inferential statistics to be used, either parametric or non-parametric statistics. Given that the sample size was below 50, the Shapiro-Wilk test was employed for continuous variables to assess normality. All selected fitness and learning attitude data not met the criteria for normality distribution for which a non-parametric test (Mann-Whitney U test) was used. Levene's test was utilised to test the homogeneity of variance for physical fitness and students' learning attitudes. Baseline data comparisons between the experimental and control groups and demographic variables were conducted using an independent-sample T-test for physical fitness and students' learning attitudes. The Kruskal-WallisH test and One-way ANOVA examined baseline data linkage between SEM and ST.

A Generalized Estimating Equation (GEE) model was employed to evaluate intervention programs' effectiveness on dependent variables. GEE extends the logistic regression model to account for clustering (McCaffrey & Bell, 2006). Typically, Cluster Randomised Controlled Trial studies utilise GEE for analysis (Barker et al., 2016; Huang et al., 2016; Liu et al., 2019). Furthermore, the effect size (d) of each analysis was automatically determined by SPSS, and Cohen's guidelines (2013) were used to interpret

effect size (d = 0.2 signifies small; d = 0.5 signifies medium, and d = 0.8 signifies large) to ascertain the effect size of the intervention. The details of the specific statistical analysis used to solve each hypothesis are shown in Table 3.12.

Table 3.12 : The Recruitment for Statistical Analysis towards Each Hypothesis

	Hypothesis	Statistical Analysis
H01-1	There were no statistically significant differences between the experimental and control groups in body composition at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.	GEE
H01-2	There were no statistically significant differences between the experimental and control groups in cardiovascular fitness at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.	GEE
H01-3	There were no statistically significant differences between the experimental and control groups in flexibility at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.	GEE
H01-4	There were no statistically significant differences between the experimental and control groups in agility at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.	GEE
H01-5	There were no statistically significant differences between the experimental and control groups in Upper body strength at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.	GEE
H01-6	There were no statistically significant differences between the experimental and control groups in Lower body explosiveness at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.	GEE
H01-7	There were no statistically significant differences between the experimental and control groups in speed at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.	GEE
H02-1	There were no statistically significant differences between the experimental and control groups in cognitive learning attitude at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.	GEE
H02-2	There were no statistically significant differences between the experimental and control groups in affective learning attitudes	GEE

Table 3.12 : Continued

	at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.	
H02-3	There were no statistically significant differences between the experimental and control groups in learning attitude at baseline, post-test1, and post-test2 among junior middle school students in Hebei, China.	GEE

Note：GEE：Generalised Estimating Equation

CHAPTER 4

RESULT

4.1 Introduction

This chapter delineates the impacts of the SEM and Standard teaching approaches in football on junior middle school students' physical fitness and learning attitudes in Hebei, China. The experimental procedures and outcomes are scrutinised following the methodology outlined in Chapter Three. Primary findings are presented and synthesised to validate the ten null hypotheses described in Chapter One.

4.2 Preliminary Examination of Statistical Hypotheses

This section primarily assesses the normality distribution and homogeneity of variance of demographic and baseline dependent variables. If the data satisfy these assumptions, parametric tests are employed. When one or both assumptions are violated, appropriate non-parametric tests are utilised based on the number of groups being compared. The Mann-Whitney U test is used for two-group comparisons, while the Kruskal-Wallis H test is applied for comparisons involving more than two groups.

The Shapiro-Wilk test indicated that all demographic variables except height exhibited non-normal distributions ($p < 0.05$), while height showed a normal distribution ($p > 0.05$). For baseline dependent variables, height, cognitive dimension score, and learning attitude score met the assumption of normality ($p > 0.05$). The homogeneity of variance test indicated that only height met

the homogeneity assumption ($p > 0.05$), whereas the cognitive dimension and learning attitude scores did not. Other baseline dependent variables showed non-normal distributions ($p < 0.05$). Please refer to Table 1-3 of Appendix K for detailed information. Therefore, a parametric test (One-way ANOVA) was conducted for height due to its normal distribution and homogeneity of variance. In contrast, the Kruskal-Wallis H test was used for other variables that did not meet these assumptions.

4.3 Descriptive Statistics of Demographic Variables and Dependent Variable at Baseline

This section discusses the intergroup demographic variables and dependent variables at baseline. Single-factor analysis of variance was conducted for height. At the same time, non-parametric tests (Kruskal-Wallis H) were utilised for family income, family cultural background, health 1, health 2, health behaviour, baseline height, baseline weight, baseline BMI-for-age, baseline sit-and-reach, baseline 30-meter sprint, baseline pull-ups, baseline YOYO test, baseline vertical jump, baseline affective dimension score, baseline cognitive dimension score, and baseline total learning attitude score. The results revealed no significant differences among intergroup demographic variables and dependent variables at baseline. Detailed data can be found in Table 4 and Table 5 of Appendix K.

4.4 Inferential Statistics

In this section, inferential statistics were employed to assess the study hypotheses using Generalized Estimating Equations (GEE), determining whether to reject or not reject the research hypotheses.

4.5 Effect of Sport Education Model on Physical Fitness

This section focuses on analysing the six hypotheses, H01-1 to H01-7. GEE techniques were employed to assess the impact of interventions on junior high school students' physical fitness and determine whether to reject the hypotheses.

4.5.1 Body Composition

The first domain of physical fitness is BMI-for-age, assessed by linking weight and height to evaluate the impact of two football teaching interventions. Furthermore, the analysis was adjusted for gender, while incorporating family income as a covariate. Descriptive data can be found in Table 4.1.

Table 4.1 : Descriptive Statistics (Mean and SE) of BMI-for-age for Control and Experimental Groups at Different Time Points

Variable	Times	Control Group LsMean (SE)	Experimental Group LsMean (SE)
BMI-for-age (kg/m^2)	Baseline	22.666 (0.794)	21.990 (0.665)
	Post-test1	22.587 (0.772)	21.995 (0.668)
	Post-test2	22.540 (0.796)	21.857 (0.669)

Based on the updated results of the GEE analysis on BMI-for-age scores, there was no statistically significant difference between the two groups in

terms of BMI-for-age (χ^2 = 0.397, df = 1, p = 0.529). Similarly, gender (χ^2 = 0.362, df = 1, p = 0.547) and family income (χ^2 = 1.910, df = 1, p = 0.167) were not significant predictors of BMI-for-age. However, time had a significant effect on BMI-for-age scores (χ^2 = 120.076, df = 2, p < 0.001), indicating that BMI-for-age scores varied significantly across different time points. Additionally, the interaction between time and group did not reach statistical significance (χ^2 = 4.978, df = 2, p = 0.083), suggesting that the changes in BMI-for-age over time did not significantly differ between the two groups (See Table 4.2).

Table 4.2 : Result of GEE on BMI-for-age Score

Variable	Source	Wald Chi-Square	df	*p*-value
BMI-for-age (kg/m^2)	Gender	0.362	1	0.547
	Family income	1.910	1	0.167
	Group	0.397	1	0.529
	Time	120.076	2	0.000
	Time * Group	4.978	2	0.083

To assess the differences in BMI-for-age among the two groups of junior high school students at different time points, Bonferroni post-hoc tests were conducted, and the results are presented in Table 4.3.

For the SEM group, there were no statistically significant differences in BMI-for-age between baseline and post-test1 (p > 0.05). However, significant differences were observed between baseline and post-test2 (p < 0.001), as well as between post-test1 and post-test2 (p < 0.05). But, for the ST group, no

significant differences were found between baseline - post-test1, baseline - post-test2, and post-test1 - post-test2 ($p > 0.05$).

Furthermore, the effect size (d) for the SEM group was greater than that of the ST group across all three time periods (baseline - post-test2 and post-test1 - post-test2), indicating that the SEM group exhibited larger changes in BMI-for-age over time compared to the ST group, but the effect size level was smaller.

Table 4.3 : Within-group Comparison of BMI-for-age at Different Time Points

Variable	Time	Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
BMI-for-age (kg/m^2)	Baseline - post-test1	ST	0.079	0.039	1	0.137	-0.016	0.173	0.015
		SEM	-0.005	0.021	1	1.000	-0.056	0.046	0.001
	Baseline - post-test2	ST	0.126	0.015	1	1.000	0.089	0.164	0.024
		SEM	0.134	0.019	1	0.001	0.087	0.180	0.028
	Post-test1 - post-test2	ST	0.048	0.037	1	0.595	-0.041	0.136	0.001
		SEM	0.139	0.017	1	0.001	0.098	0.180	0.029

To compare the BMI-for-age scores between the SEM group and the ST group at different time points (baseline, post-test1, and post-test2), post-hoc tests (Bonferroni) were conducted, as shown in Table 4.4. The results indicate that there were no statistically significant differences in BMI-for-age scores between the SEM group and the ST group at any time point: baseline ($p = 0.514$), post-test1 ($p = 0.562$), and post-test2 ($p = 0.512$).

Furthermore, effect size calculations revealed small effect sizes across all three time points: baseline (d = 0.130), post-test1 (d = 0.115), and post-test2 (d = 0.129), suggesting negligible practical differences between the groups.

In summary, these findings indicate that the SEM intervention did not produce significant changes in BMI-for-age compared to the standard teaching approach, with minimal differences observed throughout the intervention period.

Table 4.4 : Comparison of BMI-for-age between Groups at Different Time Points

Variable	Time	Between Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
BMI-for-age (kg/m^2)	Baseline	ST-SEM	0.676	1.037	1	0.514	-1.356	2.708	0.130
	Post-test1	ST-SEM	0.592	1.021	1	0.562	-1.410	2.594	0.115
	Post-test2	ST-SEM	0.683	1.041	1	0.512	-1.356	2.723	0.129

The average body composition (BMI-for-age) scores throughout the process are shown in Figure 4.1. The ST group had slightly higher baseline BMI-for-age test scores than the SEM group. However, the BMI-for-age test scores for the ST group decreased from baseline to post-test2, whereas the SEM group only showed a decrease from post-test1 to post-test2. At the end of the 12-week intervention, the BMI-for-age test scores for the SEM group remained slightly lower than those of the ST group.

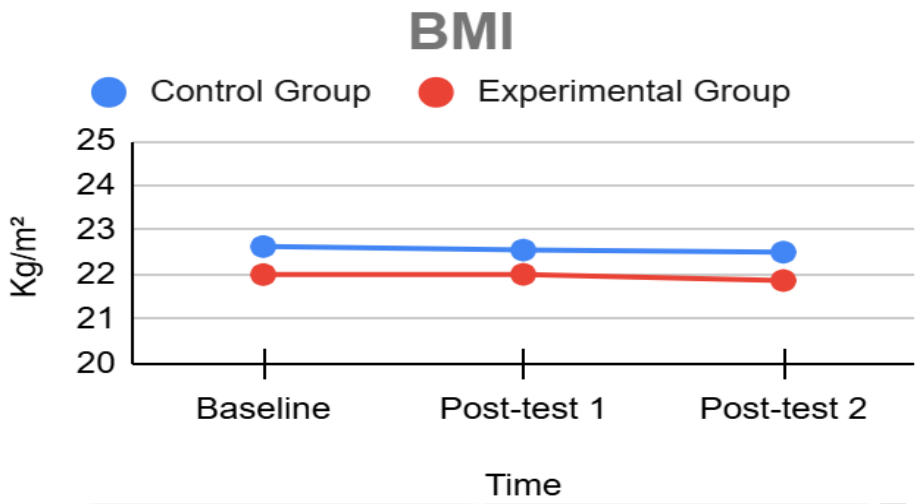


Figure 4.1 : The Mean Scores Changes of BMI-for-age across the Times for Both Groups

4.5.2 Flexibility

The second domain of physical fitness is flexibility, observed through the sit and reach test to assess the impact of two football teaching interventions. Descriptive data can be found in Table 4.5.

Table 4.5 : Descriptive Statistics (Mean and SE) of Sit and Reach for Control and Experimental Groups at Different Time Points

Variable	Times	Control Group LsMean (SE)	Experimental Group LsMean (SE)
Sit and reach (cm)	Baseline	7.433 (1.094)	8.160 (1.352)
	Post-test1	9.386 (1.304)	10.469 (1.009)
	Post-test2	11.660 (1.101)	14.969 (0.964)

Based on the GEE analysis results for the sit and reach score, the following conclusions can be drawn: There was statistically significant difference between genders regarding the sit and reach score ($\chi 2 = 22.196$, $p = 0.001$). There was no statistically significant difference between groups regarding the sit and reach score ($\chi 2 = 3.474$, $p = 0.098$). However, time significantly

affected the sit and reach score ($\chi 2 = 28.346$, $p < 0.001$). The interaction between time and group was not statistically significant ($\chi 2 = 2.274$, $p = 0.321$), indicating that the changes in sit and reach scores over time did not significantly differ between groups, and there were significant differences between the gender. (See Table 4.6)

Table 4.6 : Result of GEE on Sit and Reach Score

Variable	Source	Wald Chi-Square	df	*p* -value
Sit and reach (cm)	Gender	22.196	1	0.001
	Group	3.474	1	0.098
	Time	28.346	2	0.001
	Time * Group	2.274	2	0.321

Table 4.7 Results show that, there was no statistically significant difference between groups for boys regarding the sit and reach score ($\chi 2 = 1.069$, $p = 0.301$), and there was also no statistically significant difference between groups for girls regarding the sit and reach score ($\chi 2 = 1.069$, $p = 0.301$). This indicates that gender has no significant effect on the sit and reach score of SEM intervention students.

Table 4.7 : Analysis of Differences between Groups by Gender

Gender	Source	Wald Chi-Square	df	*p* -value
Boys	Between group	1.069	1	0.301
Girls	Between group	2.557	1	0.110

To assess the differences in sit and reach flexibility among the two groups of junior high school students at different time points, Bonferroni post-hoc tests

were conducted, and the results are presented in Table 4.8. For the SEM group, there were no statistically significant differences in sit and reach flexibility between baseline and post-test1 ($p > 0.05$). Still, significant differences between baseline and post-test2 ($p = 0.001$) were observed, indicating improved flexibility over time. However, for the ST group, there was no significant difference between baseline, post-test1, and post-test2 ($p > 0.05$), indicating that the flexibility of the ST group from baseline to post-test2 was not significantly improved. Moreover, the effect size (d) for the SEM group was notably larger than that of the ST group across all three time periods (baseline, post-test1, and post-test2), indicating that the SEM group exhibited greater improvements in sit and reach flexibility over time compared to the ST group.

Table 4.8 : Within-group Comparison of Sit and Reach at Different Time Points

Variable	Time	Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
Sit and reach (cm)	Baseline - post-test1	ST	-1.952	2.033	1	1.000	-6.819	2.914	0.238
		SEM	-2.310	1.679	1	0.507	-6.330	1.711	0.284
	Baseline - post-test2	ST	-4.226	1.790	1	0.055	-8.512	0.060	0.579
		SEM	-6.810	1.772	1	0.001	-11.052	-2.567	0.971
	Post-test1 - post-test2	ST	-2.274	1.526	1	0.409	-5.928	1.380	0.275
		SEM	-4.500	.6760	1	0.001	-6.118	-2.882	0.641

To compare the sit and reach scores between the SEM group and the ST group at different time points (baseline, post-test1, and post-test2), post-hoc tests (Bonferroni) were conducted, as shown in Table 4.9. The results indicate that there were no statistically significant differences in sit and reach

scores between the two groups at any of the time points: baseline (p = 1.000), post-test1 (p = 1.000), and post-test2 (p = 0.462).

Additionally, effect size calculations revealed small effect sizes across all three time points: baseline (d = 0.089), post-test1 (d = 0.132), and post-test2 (d = 0.466), indicating minimal differences between the SEM and ST groups.

In summary, these findings demonstrate that the SEM intervention did not result in significant improvements in flexibility compared to the standard teaching method, with only minimal differences observed over time.

Table 4.9 : Comparison of Sit and Reach between Groups at Different Time Points

Variable	Time	Between Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
Sit and reach (cm)	Baseline	ST-SEM	-.726	1.740	1	0.676	-4.136	2.684	0.089
	Post-test1	ST-SEM	-1.083	1.648	1	1.000	-4.313	2.147	0.132
	Post-test2	ST-SEM	-3.310	1.461	1	0.023	-6.172	-0.447	0.466

Figure 4.2 illustrates the average Flexibility scores (sit and reach test) for both groups throughout the process. The SEM group had slightly higher sit and reach test scores at baseline than the ST group. However, both groups showed an increase in Sit-and-reach test scores after test 1, and at the end of the 12-week intervention, the SEM group's average total Sit-and-reach test score far surpassed that of the ST group after test 2.

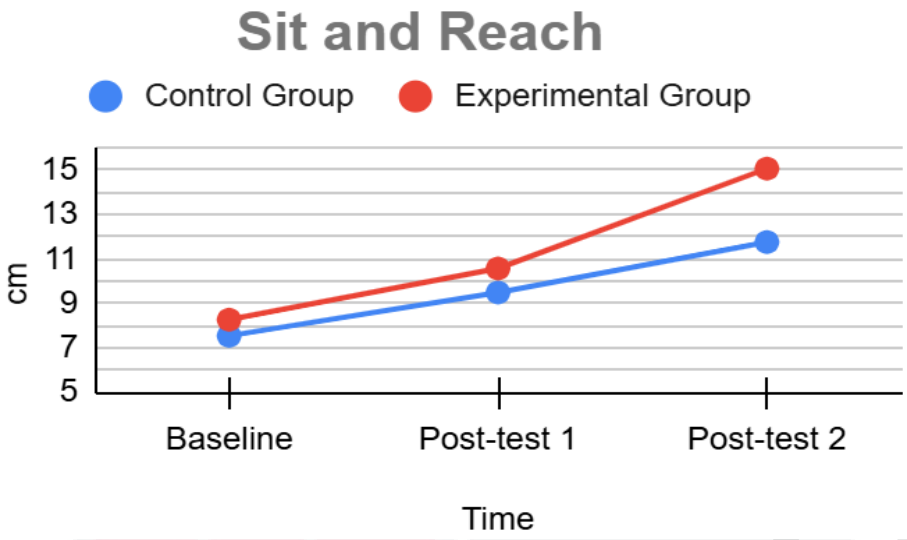


Figure 4.2 : The Mean Scores Changes of Flexibility across the Times for Both Groups

4.5.3 Agility

The third domain of physical fitness is agility, which was observed through the arrowhead agility test to assess the impact of two football teaching interventions. Descriptive data are in Table 4.10.

Table 4.10 : Descriptive Statistics (Mean and SE) of Arrowhead Agility Test for Control and Experimental Groups at Different Time Points

Variable	Times	Control Group LsMean (SE)	Experimental Group LsMean (SE)
Arrowhead Agility Test (second)	Baseline	11.156 (0.105)	11.010 (0.119)
	Post-test1	10.988 (0.134)	10.949 (0.131)
	Post-test2	10.763 (0.121)	9.966 (0.130)

Based on the GEE analysis results for the arrowhead agility test score, there was statistically significant difference between genders regarding the arrowhead agility test score ($\chi2 = 32.412$, $p = 0.001$), groups ($\chi2 = 8.062$, $p = 0.005$), Time ($\chi2 = 34.571$, $p < 0.001$), and the interaction between time and group was statistically significant ($\chi2 = 9.223$, $p = 0.010$), indicating that the changes in arrowhead agility test scores over time varied significantly

between groups, and there were significant differences between the gender. (See Table 4.11)

Table 4.11 : Result of GEE on Arrowhead Agility Test Score

Variable	Source	Wald Chi-Square	df	*p*-value
Arrowhead Agility Test (second)	Gender	32.412	1	0.001
	Group	8.062	1	0.005
	Time	34.571	2	0.001
	Time * Group	9.223	2	0.010

Table 4.12 Results show that, there was a statistically significant difference between groups for boys regarding the arrowhead agility test score ($\chi 2 = 3.996$, $p = 0.046$), and there was also a statistically significant difference between groups for girls regarding the arrowhead agility test score ($\chi 2 = 4.089$, $p = 0.043$). This indicates that gender has a significant effect on the arrowhead agility test score of SEM intervention students.

Table 4.12 : Analysis of Differences between Groups by Gender

Gender	Source	Wald Chi-Square	df	*p* -value
Boys	Between group	3.996	1	0.046
Girls	Between group	4.089	1	0.043

To assess the differences in Arrowhead Agility Test performance among the two groups of junior high school students at different time points, Bonferroni post-hoc tests were conducted, and the results are presented in Table 4.13.

There were no statistically significant differences in Arrowhead Agility Test performance for the SEM group between baseline and post-test1 ($p > 0.05$). Still, significant differences were observed between baseline and post-test2 and post-test1 - post-test2 ($p < 0.05$), indicating improved agility over time. However, for the ST group, there was no significant difference between baseline, post-test1, and post-test2 ($p > 0.05$), indicating that the Agility of the ST group was not significantly improved from baseline to post-test2.

In addition, the effect size (d) of the SEM group was significantly larger than that of the ST group at the time periods (baseline - post-test2 and post-test1-post-test2), indicating that the SEM group had a greater improvement in performance on the arrow agility test over time than the ST group, especially post-test1-post-test2.

Table 4.13 : Within-group Comparison of Arrowhead Agility Test at Different Time Points

Variable	Time	Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
Arrowhead Agility Test (second)	Baseline - post-test1	ST	0.169	0.093	1	0.212	-0.055	0.392	0.198
		SEM	0.061	0.150	1	1.000	-0.298	0.420	0.069
	Baseline - post-test2	ST	0.394	0.157	1	0.073	0.017	0.770	0.475
		SEM	1.044	0.188	1	0.001	0.593	1.495	1.227
	Post-test1 - post-test2	ST	0.225	0.177	1	0.610	-0.648	0.198	0.253
		SEM	0.983	0.184	1	0.001	0.543	1.423	1.044

To compare the Arrowhead Agility Test scores between the SEM group and the ST group at different time points (baseline, post-test1, and post-test2), post-hoc tests (Bonferroni) were conducted, as shown in Table 4.14. The

results show no statistically significant differences between the two groups at baseline (p = 0.355) and post-test1 (p = 0.838).

However, a statistically significant difference emerged at post-test2 (p = 0.001), with a mean difference of 0.796 and a large effect size (d = 0.895), indicating that the SEM group outperformed the ST group in agility at this stage. Effect size values were small at baseline (d = 0.186) and post-test1 (d = 0.041), increasing to a large effect size at post-test2 (d = 0.895).

In summary, these results suggest that while early differences in agility between groups were minimal, the SEM intervention led to a significant and substantial improvement in agility performance by the end of the intervention.

Table 4.14 : Comparison of Arrowhead Agility Test between Groups at Different Time Points

Variable	Time	Between Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
Arrowhead Agility Test (second)	Baseline	ST-SEM	0.146	0.158	1	0.355	-0.164	0.456	0.186
	Post-test1	ST-SEM	0.038	0.188	1	0.838	-0.406	0.329	0.041
	Post-test2	ST-SEM	0.796	0.178	1	0.001	0.447	1.146	0.895

Figure 4.3 displays the average completion times of agility (Arrowhead Agility test) for both groups throughout the process. Compared to the ST group, the SEM group had slightly lower average completion times in the Arrowhead Agility test at baseline. However, after test 1, the changes in Arrowhead Agility test completion times for both groups were minimal. Yet, at the end of the 12-week intervention, the SEM group exhibited a significant reduction in

average Arrowhead Agility test completion time after test 2, with a greater reduction than the ST group.

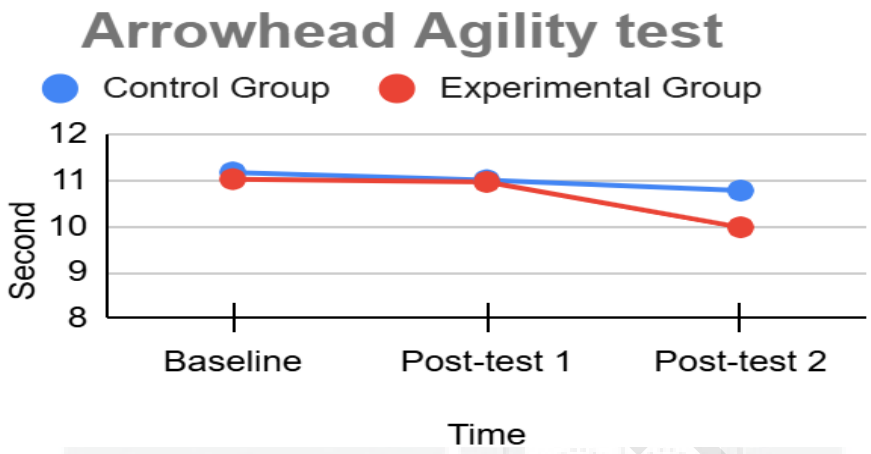


Figure 4.3 : The Mean Scores Changes of Agility across the Times for Both Groups

4.5.4 Speed

The fourth domain of physical fitness is Speed, which was observed through the 30-meter sprint test to assess the impact of two football teaching interventions. Descriptive data can be found in Table 4.15.

Table 4.15 : Descriptive Statistics (Mean and SE) of 30-meter Sprint Test for Control and Experimental Groups at Different Time Points

Variable	Times	Control Group LsMean (SE)	Experimental Group LsMean (SE)
30-meter Sprint Test (second)	Baseline	6.064 (0.100)	6.066 (0.095)
	Post-test1	6.032 (0.112)	5.969 (0.109)
	Post-test2	5.992 (0.095)	5.494 (0.087)

Based on the results of the GEE analysis on the 30-meter Sprint Test scores, there was statistically significant difference between genders regarding the sprint performance ($\chi2 = 33.105$, df = 1, $p < 0.001$), group ($\chi2 = 4.654$, df = 1,

p = 0.031), time (χ2 = 11.123, df = 2, p = 0.004), suggesting that the sprint performance varied significantly across different time points. Moreover, the interaction between time and group was statistically significant (χ2 = 6.901, df = 2, p = 0.032), indicating that the changes in sprint performance over time differed between the two groups, and there were significant differences between the gender. (See Table 16)

Table 4.16 : Result of GEE on 30-meter Sprint Test Score

Variable	Source	Wald Chi-Square	df	*p*-value
30-meter Sprint Test (second)	Gender	33.105	1	0.001
	Group	4.654	1	0.031
	Time	11.123	2	0.004
	Time * Group	6.901	2	0.032

Table 4.17 Results show that, there was no statistically significant difference between groups for boys regarding the 30-meter sprint test score (χ2 = 1.133, p = 0.287), and there was also no statistically significant difference between groups for girls regarding the 30-meter sprint test score (χ2 = 3.719, p = 0.054). This indicates that gender has no significant effect on the 30-meter sprint test score of SEM intervention students.

Table 4.17 : Analysis of Differences between Groups by Gender

Gender	Source	Wald Chi-Square	df	*p*-value
Boys	Between group	1.133	1	0.287
Girls	Between group	3.719	1	0.054

To evaluate the changes in performance in the 30-meter Sprint Test among the two groups of junior high school students at different time points, Bonferroni post-hoc tests were conducted, and the results are presented in Table 4.18.

For the SEM group, there were no statistically significant differences in 30-meter Sprint Test performance between baseline and post-test1 ($p > 0.05$). Still, significant improvements were observed between baseline and post-test2 ($p = 0.001$), indicating enhanced sprint performance over time. But for the ST group of 30 m sprint, test scores in the three periods (the Pre-test - post-test1, the Pre-test - post-test2, and post-test1 to post-test2) were not significantly different ($p > 0.05$).

Furthermore, the effect size (d) for the SEM group was notably larger than that of the ST group across all three time periods (baseline, post-test1, and post-test2), indicating that the SEM group exhibited greater improvements in the 30-meter Sprint Test performance over time compared to the ST group.

Table 4.18 : Within-group Comparison of 30-meter Sprint Test at Different Time Points

Variable	Time	Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
30-meter Sprint Test (second)	Baseline - post-test1	ST	0.032	0.070	1	1.000	-0.134	0.199	0.042
		SEM	0.097	0.136	1	1.000	-0.227	0.422	0.138
	Baseline - post-test2	ST	0.072	0.131	1	1.000	-0.242	0.386	0.113
		SEM	0.572	0.143	1	0.001	0.230	0.914	0.925
	Post-test1 - post-test2	ST	0.040	0.136	1	1.000	-0.286	0.365	0.060
		SEM	0.475	0.148	1	0.0014	0.119	0.830	0.686

To compare the 30-meter Sprint Test scores between the SEM group and the ST group at different time points (baseline, post-test1, and post-test2), post-hoc tests (Bonferroni) were conducted, as shown in Table 4.19. The results indicate no statistically significant differences between the two groups at baseline (p = 1.000) and post-test1 (p = 1.000).

However, a statistically significant difference was observed at post-test2 (p = 0.001), with a mean difference of 0.498 and a large effect size (d = 0.893), indicating that the SEM group outperformed the ST group in sprint performance at that stage. Effect sizes were negligible at baseline (d = 0.004) and post-test1 (d = 0.081), but increased markedly at post-test2 (d = 0.893).

In summary, these findings show that although no early differences in sprint performance were observed, the SEM intervention resulted in a significant and substantial improvement by the end of the intervention period.

Table 4.19 : Comparison of 30-meter Sprint Test between Groups at Different Time Points

Variable	Time	Between Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
30-meter Sprint Test (second)	Baseline	ST-SEM	-0.003	0.138	1	0.985	-0.273	0.268	0.004
	Post-test1	ST-SEM	0.063	0.156	1	0.688	-0.244	0.369	0.081
	Post-test2	ST-SEM	0.498	0.112	1	0.001	0.278	0.268	0.893

Figure 4.4 illustrates the average completion times of speed (30-meter sprint test) for both groups throughout the process. At baseline, the completion times for the 30-meter sprint test were essentially the same for both the SEM and ST groups. After post-test1, the completion time for the 30-meter sprint

test was slightly lower for the SEM group than the ST group, but the difference was insignificant. However, at the end of the 12-week intervention, the SEM group exhibited a significant reduction in average completion time for the 30-meter sprint test after test 2, with a reduction greater than that of the ST group.

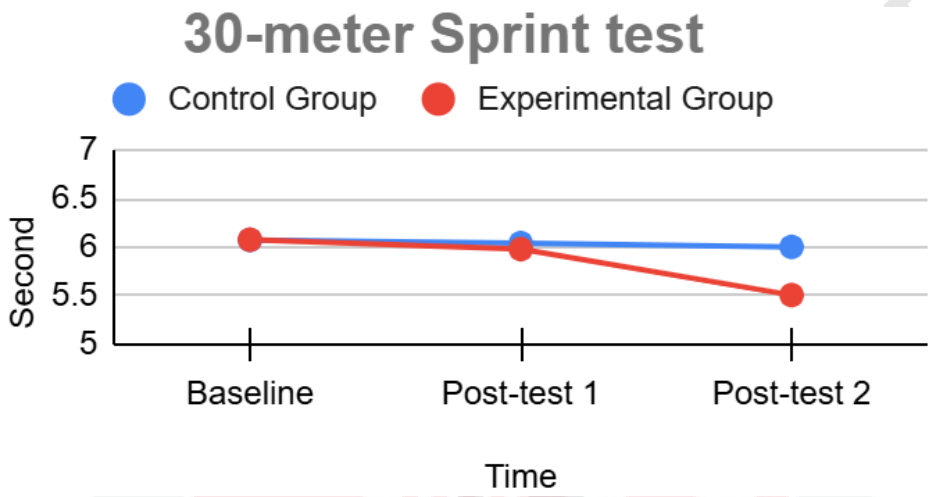


Figure 4.4 : The Mean Scores Changes of Speed across the Times for Both Groups

4.5.5 Upper Body Strength

The fifth domain of physical fitness is muscle endurance, which was observed through the pull-up test to assess the impact of two football teaching interventions. Descriptive data are in Table 4.20.

Table 4.20 : Descriptive Statistics (Mean and SE) of Pull-up Test for Control and Experimental Groups at Different Time Points

Variable	Times	Control Group LsMean (SE)	Experimental Group LsMean (SE)
Pull-up Test (Reps)	Baseline	1.043 (0.343)	1.281 (0.418)
	Post-test1	1.591 (0.431)	1.805 (0.394)
	Post-test2	2.734 (0.477)	3.686 (0.492)

Based on the results of the GEE analysis on the Pull-up Test scores, there was statistically significant difference between the genders regarding pull-up performance ($\chi 2 = 13.019$, df = 1, p = 0.001); there was no statistically significant difference between the two groups regarding pull-up performance ($\chi 2 = 0.916$, df = 1, p = 0.338). However, time significantly affected the pull-up test scores ($\chi 2 = 56.191$, df = 2, p = 0.001), indicating that the pull-up performance varied significantly across different time points. Furthermore, the interaction between time and group did not reach statistical significance ($\chi 2 = 3.145$, df = 2, p = 0.207), suggesting that the changes in pull-up performance over time did not significantly differ between the two groups, but there were significant differences between the gender (See Table 4.21).

Table 4.21 : Result of GEE on Pull-up Test Score

Variable	Source	Wald Chi-Square	df	*p*-value
Pull-up Test (Reps)	Gender	13.109	1	0.001
	Group	0.916	1	0.338
	Time	56.191	2	0.001
	Time * Group	3.145	2	0.207

Table 4.22 Results show that, there was no statistically significant difference between groups for boys regarding the pull-up test score ($\chi 2 = 0.016$, p = 0.899), and there was also no statistically significant difference between groups for girls regarding the pull-up test score ($\chi 2 = 3.780$, p = 0.052). This indicates that gender has no significant effect on the pull-up test score of SEM intervention students.

Table 4.22 : Analysis of Differences between Groups by Gender

Gender	Source	Wald Chi-Square	df	*p*-value
Boys	Between group	0.016	1	0.899
Girls	Between group	3.780	1	0.052

To evaluate changes in performance in the Pull-up Test within each group of junior high school students at different time points, P-values and effect sizes (d) were analysed, as shown in Table 4.23.

For the Pull-up Test, within the SEM group, there were no statistically significant differences between baseline and post-test1 ($p > 0.05$), and the effect size (d) was 0.193. However, significant improvements were observed between baseline and post-test2, as well as between post-test1 and post-test2 ($p < 0.05$), with effect sizes (d) of 0.788 and 0.613, respectively.

Similarly, within the ST group, there were no statistically significant differences between baseline and post-test1 ($p > 0.05$), and the effect size (d) was 0.202. Significant improvements were observed between baseline and post-test2, as well as between post-test1 and post-test2 ($p < 0.05$), with effect sizes (d) of 0.586 and 0.361, respectively.

Therefore, while the effect size for the ST group was slightly higher than that of the SEM group in the Baseline - post-test1 time period, the SEM group exhibited larger effect sizes in the Baseline - post-test2 and post-test1 - post-test2 time periods, indicating greater improvements in Pull-up Test performance over time compared to the ST group.

Table 4.23 : Within-group Comparison of Pull-up Test at Different Time Points

Variable	Time	Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
Pull-up Test (Reps)	Baseline - post-test1	ST	-0.548	0.274	1	0.137	-1.204	0.108	0.202
		SEM	-0.524	0.546	1	1.000	-1.832	0.784	0.193
	Baseline - post-test2	ST	-1.691	0.341	1	0.001	-2.506	-0.874	0.586
		SEM	-2.405	0.651	1	0.001	-3.964	-0.846	0.788
	Post-test1 - post-test2	ST	-1.143	0.212	1	0.001	-1.650	-0.636	0.361
		SEM	-1.881	0.358	1	0.001	-2.739	-1.023	0.613

To compare the Pull-up Test scores between the SEM group and the ST group at different time points (baseline, post-test1, and post-test2), post-hoc tests (Bonferroni) were conducted, as shown in Table 4.24. The results indicate that there were no statistically significant differences between the two groups at baseline ($p = 0.656$), post-test1 ($p = 0.711$), and post-test2 ($p = 0.161$).

Additionally, the effect sizes were small across all time points: baseline ($d = 0.094$), post-test1 ($d = 0.075$), and post-test2 ($d = 0.285$), suggesting minimal differences in upper body strength between the SEM and ST groups.

In summary, these results show that the SEM intervention did not lead to significant improvements in upper body strength compared to the standard teaching approach.

Table 4.24 : Comparison of Pull-up Test between Groups at Different Time Points

Variable	Time	Between Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
Pull-up Test (Reps)	Baseline	ST-SEM	-0.238	0.534	1	0.656	-1.285	0.809	0.094
	Post-test1	ST-SEM	-0.214	0.577	1	0.711	-1.346	0.917	0.075
	Post-test2	ST-SEM	-0.952	0.679	1	0.161	-2.284	0.379	0.285

Figure 4.5 presents the average muscle endurance (pull-up) scores for both groups throughout the process. At baseline, the SEM group's average pull-up score was slightly higher than the ST group's, but the difference was insignificant. After post-test1, the SEM and ST groups showed some improvement in average pull-up scores, with SEM still slightly higher than ST, though the difference was minor. However, at the end of the 12-week intervention, both SEM and ST groups exhibited improvements in average pull-up scores after test 2. Still, the improvement was more significant for the SEM group.

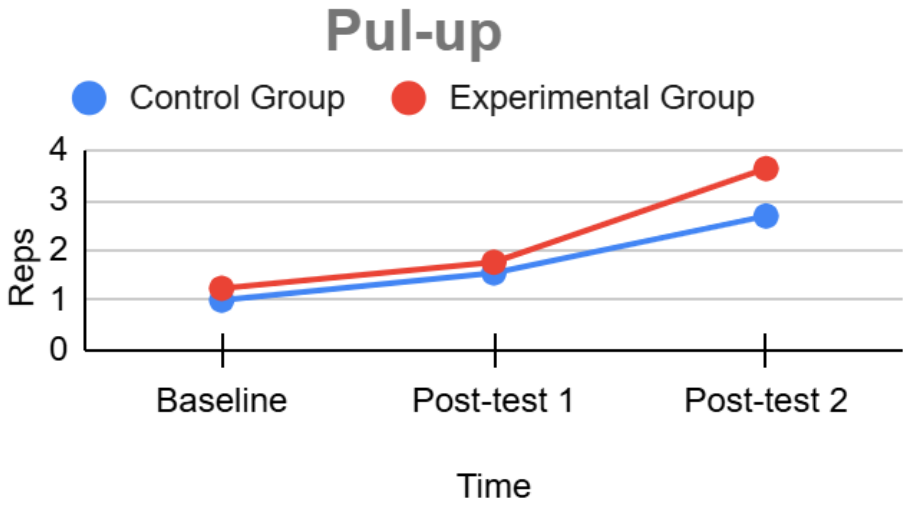


Figure 4.5 : The Mean Scores Changes of Muscle Endurance across the Times for Both Groups

4.5.6 Lower Body Explosiveness

The sixth domain of physical fitness is muscle endurance, which was observed through the pull-up test to assess the impact of two football teaching interventions. Descriptive data are in Table 4.25.

Table 4.25 : Descriptive Statistics (Mean and SE) of Vertical Jump Test for Control and Experimental Groups at Different Time Points

Variable	Times	Control Group LsMean (SE)	Experimental Group LsMean (SE)
Vertical Jump Test (cm)	Baseline	29.115 (1.026)	30.973 (1.148)
	Post-test1	30.377 (0.045)	32.925 (0.975)
	Post-test2	32.139 (1.200)	39.734 (1.507)

Based on the results of the GEE analysis on the Vertical Jump Test scores, there was a statistically significant difference between the two groups regarding vertical jump performance ($\chi 2 = 10.041$, df = 1, p = 0.002). Additionally, time significantly affected the vertical jump test scores ($\chi 2 = 43.593$, df = 2, p = 0.001), indicating that the vertical jump performance varied significantly across different time points. Moreover, the interaction between time and group was statistically significant ($\chi 2 = 12.430$, df = 2, p = 0.002). suggesting that the changes in vertical jump performance over time differed between the two groups. In addition, there was statistically significant difference between the genders regarding Vertical Jump performance ($\chi 2 = 13.552$, df = 1, p = 0.001). (See Table 4.26).

Table 4.26 : Result of GEE on Vertical Jump Test Score

Variable	Source	Wald Chi-Square	df	*p*-value
Vertical Jump Test (cm)	Gender	13.552	1	0.001
	Group	10.041	1	0.002
	Time	43.593	2	0.001
	Time * Group	12.430	2	0.002

Table 4.27 Results show that, there was a statistically significant difference between groups for boys regarding the vertical jump test score ($\chi 2 = 5.658$, $p = 0.017$), and there was also a statistically significant difference between groups for girls regarding the vertical jump test score ($\chi 2 = 4.417$, $p = 0.036$). This indicates that gender has a significant effect on the vertical jump test score of SEM intervention students.

Table 4.27 : Analysis of Differences between Groups by Gender

Gender	Source	Wald Chi-Square	df	*p*-value
Boys	Between group	5.658	1	0.017
Girls	Between group	4.417	1	0.036

For the Vertical Jump Test, within the SEM group, there were no statistically significant differences between baseline and post-test1 ($p > 0.05$), with an effect size (d) of 0.268. However, significant improvements were observed between baseline and post-test2, as well as between post-test1 and post-test2 ($p < 0.05$), with effect sizes (d) of 0.945 and 0.764, respectively.

Similarly, within the ST group, there were no statistically significant differences between baseline and post-test1 ($p > 0.05$), with an effect size (d)

of 0.190. However, significant improvements were observed between baseline and post-test2, as well as between post-test1 and post-test2 ($p < 0.05$), with effect sizes (d) of 0.400 and 0.245, respectively.

Therefore, both the ST and SEM groups exhibited significant improvements in Vertical Jump Test performance over time, with the SEM group demonstrating larger effect sizes in all periods, indicating greater improvements in vertical jump performance compared to the ST group. (See Table 4.28).

Table 4.28 : Within-group Comparison of Vertical Jump Test at Different Time Points

Variable	Time	Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
Vertical Jump Test (cm)	Baseline - post-test1	ST	-1.262	0.814	1	0.363	-3.210	0.686	0.190
		SEM	-1.952	1.497	1	0.577	-5.537	1.632	0.268
	Baseline - post-test2	ST	-3.024	0.886	1	0.002	-5.145	-0.903	0.400
		SEM	-8.762	1.722	1	0.001	-12.884	-4.640	0.945
	Post-test1 - post-test2	ST	-1.762	0.701	1	0.036	-3.439	-0.085	0.245
		SEM	-6.810	1.405	1	0.001	-10.173	-3.447	0.764

To compare the Vertical Jump Test scores between the SEM group and the ST group at different time points (baseline, post-test1, and post-test2), post-hoc tests (Bonferroni) were conducted, as shown in Table 4.29. The results indicate no statistically significant differences between the two groups at baseline ($p = 0.227$) and post-test1 ($p = 0.055$).

However, a statistically significant difference was observed at post-test2 ($p < 0.001$), with a mean difference of -7.595 and a large effect size ($d = 0.807$),

showing that the SEM group outperformed the ST group in lower body explosiveness at that stage. The effect size was small at baseline (d = 0.251), moderate at post-test1 (d = 0.391), and large at post-test2 (d = 0.807), indicating a progressive improvement.

In summary, these findings demonstrate that the SEM intervention significantly enhanced lower body power compared to the standard teaching approach by the end of the intervention.

Table 4.29 : Comparison of Vertical Jump Test between Groups at Different Time Points

Variable	Time	Between Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
Vertical Jump Test (cm)	Baseline	ST-SEM	-1.857	1.536	1	0.227	-4.868	1.153	0.251
	Post-test1	ST-SEM	-2.548	1.327	1	0.055	-5.149	0.535	0.391
	Post-test2	ST-SEM	-7.595	1.923	1	0.000	-11.365	-3.826	0.807

Figure 4.6 presents the average power (vertical jump test) scores for both groups throughout the process. At baseline, the SEM group's average score for the Vertical jump test was slightly higher than that of the ST group, but the difference was insignificant. After post-test1, the SEM and ST groups showed some improvement in average vertical jump test scores, with SEM still slightly higher than ST, though the difference was minor. However, at the end of the 12-week intervention, both SEM and ST groups exhibited improvements in average vertical jump test scores at post-test2, but the improvement was more significant for the SEM group.

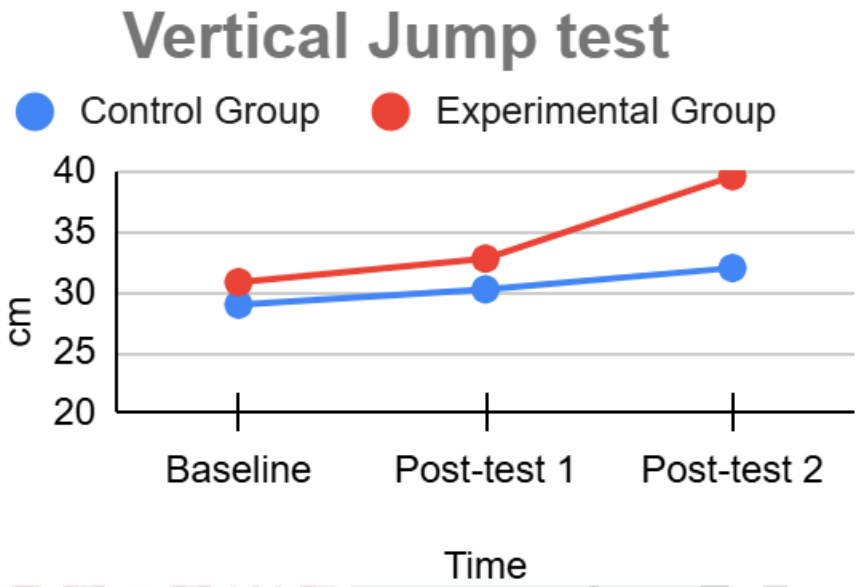


Figure 4.6 : The Mean Scores Changes of Power across the Times for Both Groups

4.5.7 Cardiovascular Fitness

The seventh domain of physical fitness is cardiovascular fitness, which was observed through the YO-YO test to assess the impact of two football teaching interventions. Descriptive data are in Table 4.30.

Table 4.30 : Descriptive Statistics (Mean and SE) of YO-YO Test for NT and SEM Groups at Different Time Points

Variable	Times	ST group LsMean (SE)	SEM Group LsMean (SE)
YO-YO Test (meter)	Baseline	344.799 (31.502)	357.180 (23.557)
	Post-test1	413.847 (30.188)	441.466 (35.988)
	Post-test2	470.038 (49.848)	698.609 (46.192)

The results of the YO-YO Test (cardiovascular fitness) indicate a statistically significant difference between genders ($\chi 2 = 15.518$, $p = 0.001$), but no statistically significant difference between two groups ($\chi 2 = 4.140$, $p = 0.042$). Furthermore, there was a significant effect of time on the YO-YO Test scores ($\chi 2 = 80.062$, $p < 0.001$). The interaction between time and group was

statistically significant ($\chi 2 = 15.613$, $p < 0.001$), suggesting that the changes in cardiovascular fitness over time differed between the two groups (baseline, post-test1, and post-test2), and there were significant differences between the sexes. Table 4.31 provides additional details.

Table 4.31 : Result of GEE on YO-YO Test Score

Variable	Source	Wald Chi-Square	df	*p*-value
YO-YO Test (meter)	Gender	15.518	1	0.001
	Group	4.140	1	0.042
	Time	80.062	2	0.001
	Time * Group	15.613	2	0.001

Table 4.32 Results show that, there was no statistically significant difference between groups for boys regarding the YO-YO test score ($\chi 2 = 0.464$, $p = 0.496$), but there was no statistically significant difference between groups for girls regarding the YO-YO test score ($\chi 2 = 3.092$, $p = 0.064$). This indicates that gender has no significant effect on the YO-YO test score of SEM intervention students.

Table 4.32 : Analysis of Differences between Groups by Gender

Gender	Source	Wald Chi-Square	df	*p*-value
Boys	Between group	0.464	1	0.496
Girls	Between group	3.092	1	0.064

Bonferroni post-hoc tests were employed to evaluate the differences in cardiovascular fitness among the two groups of junior high school students at three different time points, as shown in Table 4.33. The Bonferroni tests

revealed that the SEM group exhibit a statistically significant difference in YO-YO Test scores between baseline and post-test1, post-test1 and post-test2 ($p < 0.05$). Conversely, in the ST group, YO-YO Test scores showed significant differences between baseline and post-test1 ($p < 0.05$), as well as between baseline and post-test2 ($p < 0.05$).

In contrast, no significant difference was found between post-test1 and 2 ($p > 0.05$). The effect size (d) for the SEM group was greater than that of the ST group across all three time periods (baseline, post-test1, and post-test2), indicating that the SEM group exhibited greater changes in cardiovascular fitness over time compared to the ST group. Table 4.33 provides additional details.

Table 4.33 : Within-group Comparison of YO-YO Test at Different Time Points

Variable	Time	Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
YO-YO Test (meter)	Baseline - post-test1	ST	-69.048	11.629	1	0.001	-96.888	-41.208	0.323
		SEM	-84.286	30.732	1	0.018	-157.858	-10.714	0.415
	Baseline - post-test2	ST	-125.238	27.644	1	0.001	-191.418	-59.059	0.425
		SEM	-341.429	47.578	1	0.001	-455.329	-227.528	1.372
	Post-test1 - post-test2	ST	-56.191	30.892	1	0.207	-130.145	17.764	0.192
		SEM	-257.143	48.471	1	0.001	-373.182	-141.104	0.889

To compare the YO-YO Test scores of the SEM group and ST group at different time points (baseline, post-test1, and post-test2), post-hoc tests (Bonferroni) were conducted, as shown in Table 4.34. The results indicate no statistically significant difference in YO-YO Test scores between the SEM and ST groups before testing ($p = 0.751$) and after post-test1 ($p = 0.554$) following

adjustments. However, a statistically significant difference was observed between the SEM and ST groups in YO-YO Test scores at post-test2 (p = 0.001). Calculations of the effect size for the SEM and ST groups at three different time points revealed small effect sizes for baseline (d = 0.068) and post-test1 (d = 0.119). In contrast, post-test2 exhibited an effect size (d = 0.673) that was more than moderate level on participants' YO-YO Test scores. (See Table 4.34)

Table 4.34 : Comparison of YO-YO Test between Groups at Different Time Points

Variable	Time	Between Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
YO-YO Test (meter)	Baseline	ST-SEM	-12.381	38.952	1	0.751	-88.726	63.965	0.068
	Post-test1	ST-SEM	-27.619	46.710	1	0.554	-119.169	63.931	0.119
	Post-test2	ST-SEM	-228.571	67.671	1	0.001	-361.204	-95.939	0.673

Figure 4.7 illustrates the mean scores of cardiovascular fitness (YO-YO Test) over the entire duration. The SEM group exhibited slightly higher YO-YO Test scores than the ST group at baseline. However, both groups showed an increase in YO-YO Test scores at post-test1, and by the end of the 12-week intervention, the SEM group's average total YO-YO Test score far surpassed that of the ST group at post-test2.

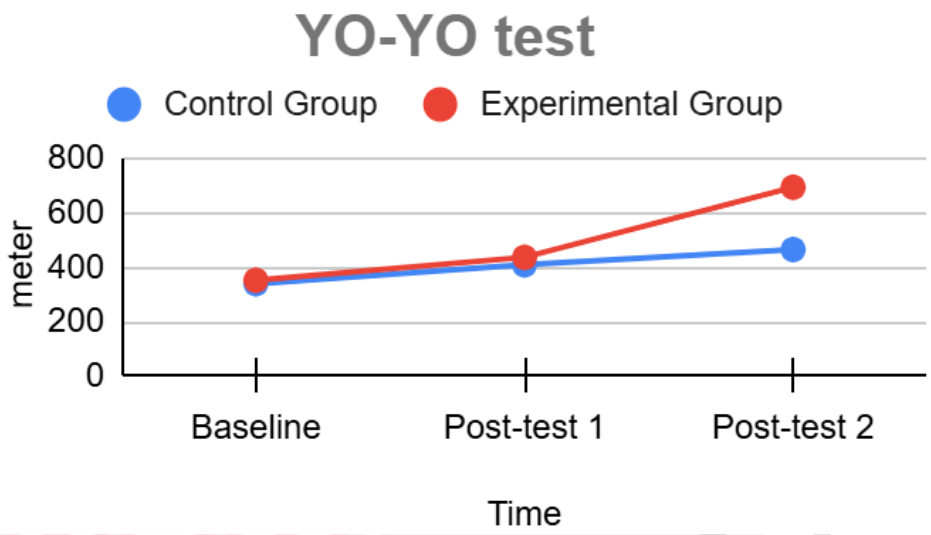


Figure 4.7 : The Mean Scores Changes of Cardiovascular Fitness across the Times for Both Groups

4.6 Effect of Sport Education Model on Learning Attitude

4.6.1 Cognitive Learning Attitude

The first domain of learning attitude is cognitive learning attitude, which is assessed through the SPEA test to evaluate the impact of two football teaching interventions. Descriptive data can be found in Table 4.35.

Table 4.35 : Descriptive Statistics (Mean and SE) of Cognitive Dimension Score for Control and Experimental Groups at Different Time Points

Variable	Times	Control Group LsMean (SE)	Experimental Group LsMean (SE)
Cognitive Dimension Score (points)	Baseline	38.620 (0. 679)	38.000 (0.570)
	Post-test1	37.640 (0.521)	41.640 (0.428)
	Post-test2	38.620 (0.363)	44.120 (0.403)

Several significant findings emerge based on the results of the GEE analysis on the Cognitive Dimension Score. First, the two groups had a statistically

significant difference regarding cognitive dimension scores ($\chi 2 = 34.134$, df = 1, p = 0.001). This suggests that there were disparities in cognitive performance between the groups. Additionally, time significantly affected cognitive dimension scores ($\chi 2 = 32.845$, df = 2, p = 0.001), indicating that cognitive performance varied significantly across different time points. Furthermore, the interaction between time and group was also statistically significant ($\chi 2 = 25.872$, df = 2, p = 0.001), suggesting that the changes in cognitive performance over time differed between the two groups. (See Table 4.36)

Table 4.36 : Result of GEE on Cognitive Dimension Score

Variable	Source	Wald Chi-Square	df	*p*-value
Cognitive Dimension Score (points)	Group	34.134	1	0.001
	Time	32.845	2	0.001
	Time * Group	25.872	2	0.001

For the Cognitive Dimension Score, within the SEM group, there were statistically significant differences between baseline and post-test1, baseline and post-test2, and post-test1 and post-test2 ($p < 0.001$), with the large effect sizes (d) of 1.101, 1.889, and 0.908.

Within the ST group, there were no statistically significant differences between baseline and post-test1 ($p > 0.05$), with a small effect size (d) of 0.192, nor between baseline and post-test2 ($p > 0.05$), with a very small effect size (d) of 0.001; additionally, there were no significant differences between post-test1 and post-test2 ($p > 0.05$), with an effect size (d) of 0.331.

Therefore, the SEM group demonstrated larger effect sizes than the ST group in all periods, indicating greater improvements in cognitive performance. (See Table 4.37)

Table 4.37 : Within-group Comparison of Cognitive Dimension Score at Different Time Points

Variable	Time	Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
Cognitive Dimension Score (points)	Baseline - post-test1	ST	0.980	1.025	1	1.000	-2.030	3.980	0.192
		SEM	-3.640	0.706	1	0.001	-5.720	-1.570	1.101
	Baseline - post-test2	ST	0.000	1.019	1	1.000	-2.990	2.990	0.001
		SEM	-6.120	0.642	1	0.001	-8.000	-4.240	1.889
	Post-test1 - post-test2	ST	-0.980	0.586	1	1.000	-2.700	0.740	0.331
		SEM	-2.480	0.614	1	0.001	-4.280	-0.670	0.908

To compare the Cognitive Dimension Score between the SEM group and ST group at different time points (baseline, post-test1, and post-test2), post-hoc tests (Bonferroni) were conducted, as shown in Table 4.38.

The results indicate no statistically significant differences in the baseline Cognitive Dimension Score between the SEM and ST groups (p = 1.000). However, significant differences were observed in the Cognitive Dimension Score between the SEM and ST groups at post-test1 (p = 0.001) and post-test2 (p = 0.001).

At post-test1, the SEM group exhibited a mean difference of -4.000 with an effect size (d) of 1.278, while at post-test2, the SEM group showed a mean difference of -5.500 with an effect size (d) of 2.188. These effect sizes

indicate large differences in Cognitive Dimension Scores between the SEM and ST groups at post-tests 1 and post-test2.

Overall, this analysis suggests that the SEM group demonstrated significantly lower Cognitive Dimension Scores than the ST group at post-test1 and post-test2, with substantial effect sizes indicating notable differences in performance between the two groups at those time points.

Table 4.38 : Comparison of Cognitive Dimension Score between Groups at Different Time Points

Variable	Time	Between Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
Cognitive Dimension Score (points)	Baseline	ST-SEM	0.620	1.123	1	1.000	-2.680	3.910	0.119
	Post-test1	ST-SEM	-4.000	0.675	1	0.001	-5.980	-2.020	1.278
	Post-test2	ST-SEM	-5.500	0.542	1	0.001	-7.090	-3.910	2.188

Figure 4.8 presents the average cognitive dimension scores for both groups throughout the process. At baseline, the ST group's scores were slightly higher than the SEM group's, but the difference was insignificant. After post-test1, there was a slight decrease in cognitive dimension scores for the ST group, while the SEM group showed a significant improvement. At post-test2, the cognitive dimension scores for the ST group students increased slightly but not significantly, whereas the SEM group's scores increased more notably compared to the ST group.

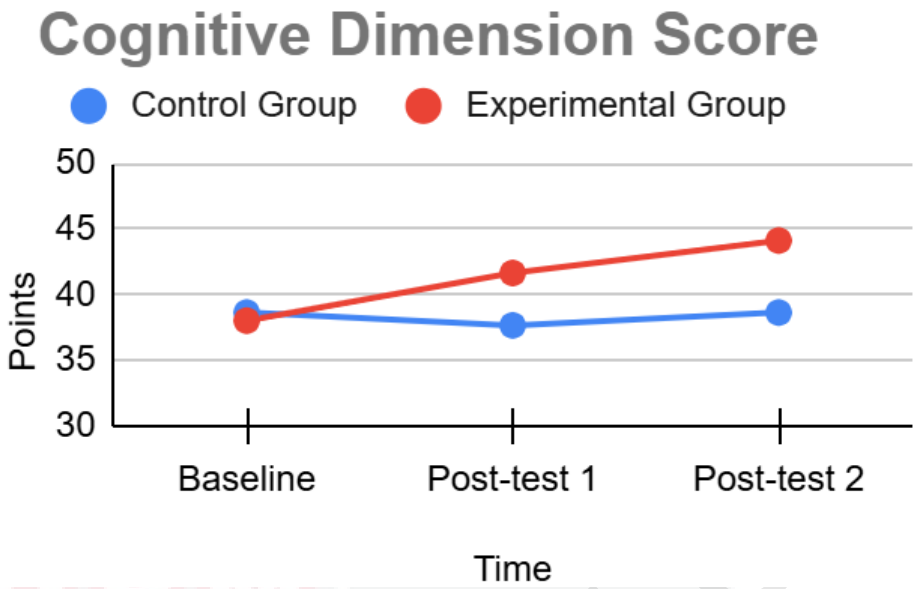


Figure 4.8 : The Mean Score Changes of Cognitive Learning Attitude across the Times for Both Groups

4.6.2 Affective Learning Attitude

The second domain of learning attitude is affective learning attitude, which is assessed through the SPEA test to evaluate the impact of two football teaching interventions. Descriptive data can be found in Table 4.39.

Table 4.39 : Descriptive Statistics (Mean and SE) of Affective Dimension Score for Control and Experimental Groups at Different Time Points

Variable	Times	Control Group LsMean (SE)	Experimental Group LsMean (SE)
Affective Dimension Score (points)	Baseline	35.810 (0.849)	36.140 (0.704)
	Post-test1	36.170 (0.414)	40.710 (0.374)
	Post-test2	37.48 (0.670)	43.860 (0.401)

Based on the results of the GEE analysis on the Affective Dimension Score, several key findings emerge. First, the two groups had a statistically significant difference regarding affective dimension scores ($\chi 2 = 63.112$, df = 1, $p = 0.001$). This indicates that there were notable distinctions in affective

responses between the groups. Additionally, time significantly affected affective dimension scores (χ2 = 51.095, df = 2, p = 0.001), suggesting that affective responses varied significantly across different time points. Furthermore, the interaction between time and group was also statistically significant (χ2 = 20.281, df = 2, p = 0.001), indicating that the changes in affective responses over time differed between the two groups (see Table 40).

Table 4.40 : Result of GEE on Affective Dimension Score

Variable	Source	Wald Chi-Square	df	*p*-value
Affective Dimension Score (points)	Group	63.112	1	0.001
	Time	51.095	2	0.001
	Time * Group	20.281	2	0.001

For the Affective Dimension Score, within the SEM group, statistically significant differences were found between baseline and post-test1 ($p < 0.001$), with a large effect size (d) of 1.237. Similarly, significant improvements were observed between baseline and post-test2 ($p < 0.001$), with a very large effect size (d) of 2.054. Additionally, significant differences were noted between post-test1 and post-test2 ($p< 0.001$), with an effect size (d) of 1.237.

Within the ST group, there were no statistically significant differences between baseline and post-test1 ($p > 0.05$), with a small effect size (d) of 0.082. Likewise, there were no significant differences between baseline and post-test2 ($p > 0.05$), with a moderate effect size (d) of 0.332. Additionally,

there were no significant differences between post-test1 and post-test2 ($p > 0.05$), with an effect size (d) of 0.358. (See Table 4.41)

Therefore, the SEM group exhibited significant improvements in Affective Dimension Score over time, while the ST group did not show significant changes. Additionally, the SEM group demonstrated larger effect sizes than the ST group in all periods, indicating greater improvements in affective responses.

Table 4.41 : Within-group Comparison of Affective Dimension Score at Different Time Points

Variable	Time	Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
Affective Dimension Score (points)	Baseline - post-test1	ST	-0.360	1.011	1	1.000	-3.320	2.610	0.082
		SEM	-4.570	0.764	1	0.001	-6.810	-2.330	1.237
	Baseline - post-test2	ST	-1.670	1.105	1	1.000	-4.910	1.580	0.332
		SEM	-7.710	0.764	1	0.001	-9.960	-5.470	2.054
	Post-test1 - post-test2	ST	-1.310	0.839	1	1.000	-3.770	1.150	0.358
		SEM	-3.140	0.531	1	0.001	-4.700	-1.580	1.237

To compare the Affective Dimension Score between the SEM group and ST group at different time points (baseline, post-test1, and post-test2), post-hoc tests (Bonferroni) were conducted, as shown in Table 4.42.

The results indicate no statistically significant differences in Affective Dimension Score between the SEM and ST groups at baseline ($p = 1.000$). However, significant differences were observed between the SEM and ST groups in the Affective Dimension Score at post-test1 ($p = 0.001$) and post-test2 ($p = 0.001$).

At post-test1, the SEM group exhibited a mean difference of -4.550 with an effect size (d) of 1.758, while at post-test2, the SEM group showed a mean difference of -6.380 with an effect size (d) of 1.762. These effect sizes indicate large differences in the Affective Dimension Score between the SEM and ST groups at both post-tests 1 and post-test2.

Overall, this analysis suggests that the SEM group demonstrated significantly lower Affective Dimension Scores than the ST group at post-test1 and post-test2, with substantial effect sizes indicating notable differences in performance between the two groups at those time points.

Table 4.42 : Comparison of Affective Dimension Score between Groups at Different Time Points

Variable	Time	Between Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
Affective Dimension Score (points)	Baseline	ST-SEM	-0.330	1.102	1	1.000	-3.570	2.900	0.065
	Post-test1	ST-SEM	-4.550	0.558	1	0.001	-6.180	-2.910	1.758
	Post-test2	ST-SEM	-6.380	0.781	1	0.001	-8.670	-4.090	1.762

Figure 4.9 illustrates the average scores of the affective dimension for both groups throughout the process. At baseline, the SEM group's scores were slightly higher than the ST group, but the difference was insignificant. After post-test1, there was no significant change in affective dimension scores for the ST group, while the SEM group showed a notable improvement. After post-test2, the affective dimension scores for the ST group students increased slightly but not significantly compared to baseline. In contrast, the SEM group's scores significantly increased more than the ST group.

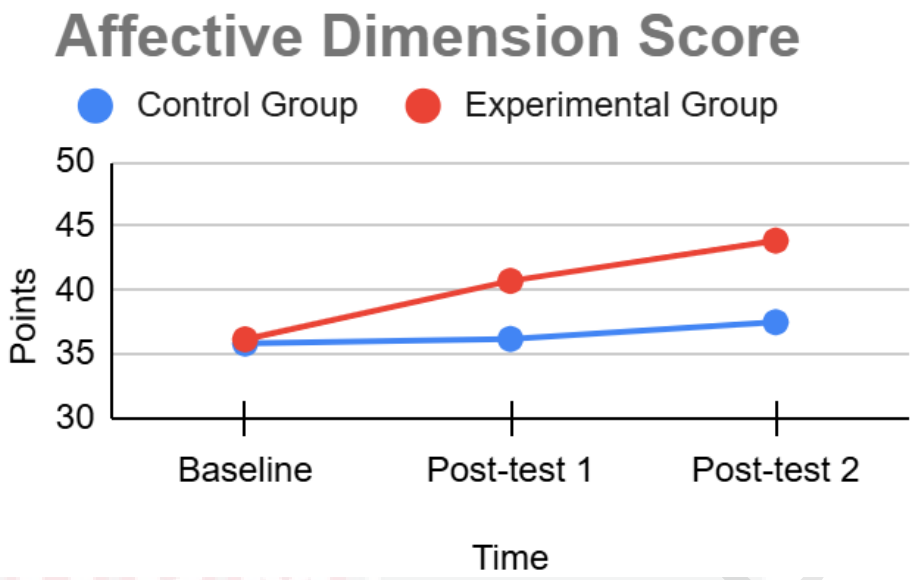


Figure 4.9 : The Mean Score Changes of Affective Learning Attitude across the Times for Both Groups

4.6.3 Holistic Learning Attitude

The Holistic learning attitude score was assessed using the SPEA test to evaluate the impact of two football teaching interventions. Descriptive data are presented in Table 4.43.

Table 4.43 : Descriptive Statistics (Mean and SE) of Holistic Learning Attitude Score for Control and Experimental Groups at Different Time Points

Variable	Times	Control Group LsMean (SE)	Experimental Group LsMean (SE)
Learning Attitude Score (points)	Baseline	74.430 (1.710)	74.140 (1.152)
	Post-test1	73.810 (0.746)	82.360 (0.612)
	Post-test2	76.100 (0.825)	87.980 (0.695)

The analysis of the Holistic Learning Attitude Score, based on the GEE results, reveals several significant findings: There was a statistically significant difference between the two groups regarding holistic learning

attitude scores (χ2 = 61.457, df = 1, p = 0.001), indicating notable disparities in holistic learning attitudes. Significant variation was observed in holistic learning attitude scores across different time points (χ2 = 56.337, df = 2, p = 0.001), suggesting that holistic learning attitudes fluctuated significantly over time. Additionally, the interaction between time and group was statistically significant (χ2 = 29.438, df = 2, p = 0.001), indicating differential changes in holistic learning attitudes over time between the two groups. (See Table 4.44)

Table 4.44 : Result of GEE on Holistic Learning Attitude Score

Variable	Source	Wald Chi-Square	df	*p*-value
Holistic Learning Attitude Score (points)	Group	61.457	1	0.001
	Time	56.337	2	0.001
	Time * Group	29.438	2	0.001

For the Holistic Learning Attitude Score, within the SEM group, there were statistically significant differences between baseline and post-test1 (p = 0.001), with a large effect size (d) of 1.357. Similarly, significant differences were observed between baseline and post-test2 (p = 0.001), with a very large effect size (d) of 2.216. Additionally, there were significant differences between post-tests 1 and 2 (p = 0.001), with an effect size (d) of 1.307.

Within the ST group, there were no statistically significant differences between baseline and post-test1 (p = 1.000), with a small effect size (d) of 0.072. Likewise, there were no significant differences between baseline and post-test2 (p = 1.000), with a small effect size (d) of 0.189. Additionally, there

were no significant differences between post-test1 and 2 (p = 0.775), with an effect size (d) of 0.443. (See Table 4.45)

Therefore, the SEM group improved its Holistic Learning Attitude Score over time, while the ST group did not change significantly. Additionally, the SEM group demonstrated larger effect sizes than the ST group in all periods, indicating greater improvement in holistic learning attitudes.

Table 4.45 : Within-group Comparison of Holistic Learning Attitude Score at Different Time Points

Variable	Time	Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
Learning Attitude Score (points)	Baseline - post-test1	ST	0.620	1.853	1	1.000	-4.820	6.060	0.072
		SEM	-8.210	1.266	1	0.001	-11.930	-4.500	1.357
	Baseline - post-test2	ST	-1.670	1.885	1	1.000	-7.200	3.870	0.189
		SEM	-13.830	1.222	1	0.001	-17.420	-10.250	2.216
	Post-test1 - post-test2	ST	-2.290	1.175	1	0.775	-5.730	1.160	0.443
		SEM	-5.620	0.990	1	0.001	-8.520	-2.710	1.307

To compare the Holistic Learning Attitude Score between the SEM group and ST group at different time points (baseline, post-test1, and post-test2), post-hoc tests (Bonferroni) were conducted, as shown in Table 4.46.

The results indicate no statistically significant differences in Holistic Learning Attitude Scores between the SEM and ST groups at baseline (p = 1.000). However, significant differences were observed between the SEM and ST groups in Holistic Learning Attitude Scores at post-test1 (p = 0.001) and post-test2 (p = 0.001).

At post-test1, the SEM group exhibited a mean difference of -8.550 with an effect size (d) of 0.443, while at post-test2, the SEM group showed a mean difference of -11.880 with an effect size (d) of 2.375. These effect sizes indicate moderate to large differences in Holistic Learning Attitude Scores between the SEM and ST groups at post-test1 and post-test2.

Overall, this analysis suggests that the SEM group demonstrated significantly lower Holistic Learning Attitude Scores than the ST group at post-test1 and post-test2, with substantial effect sizes indicating notable differences in performance between the two groups at those time points.

Table 4.46 : Comparison of Holistic Learning Attitude Score between Groups at Different Time Points

Variable	Time	Between Group	Mean Difference (I-J)	SE	df	*p*-value	95% CI		Effect Size (d)
							Lower	Upper	
Learning Attitude Score (points)	Baseline	ST-SEM	0.290	2.062	1	1.000	-5.770	6.340	0.030
	Post-test1	ST-SEM	-8.550	0.965	1	0.001	-11.380	-5.710	0.443
	Post-test2	ST-SEM	-11.880	1.079	1	0.001	-15.050	-8.710	2.375

Figure 4.10 presents the average scores of the holistic learning attitude for both groups throughout the process. At baseline, the scores for the SEM group were almost identical to those of the ST group. After post-test1, there was a slight decrease in the holistic learning attitude score for the ST group, but the decrease was insignificant. In contrast, the SEM group showed a noticeable improvement. After post-test2, compared to the ST group, the holistic learning attitude score for ST group students increased slightly but not significantly, whereas the SEM group's score exhibited a more significant increase.

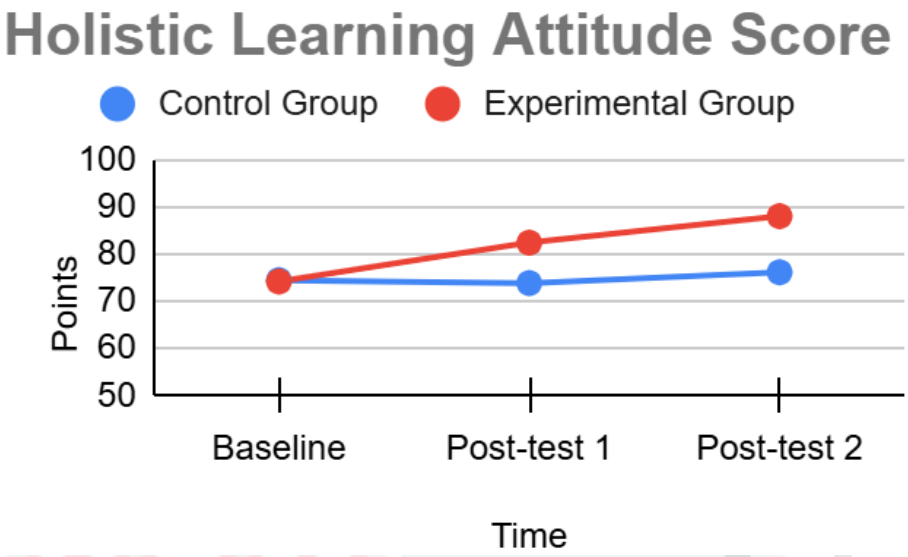


Figure 4.10 : The Mean Scores Changes of Holistic Learning Attitude across the Times for Both Groups

4.7 Conclusion

The research results indicate that compared to the ST group, participants in the SEM group (middle school students aged 13-15) showed statistically significant improvements in physical fitness and learning attitude variables. Between-group effects revealed statistically significant differences in the four physical fitness variables and all learning attitude variables between the SEM and ST groups ($p<0.05$). Specifically, all physical fitness and learning attitude variables in the SEM group from baseline to post-test2 showed statistical significance; from baseline to post-test1, cognitive dimension score, affective dimension score, and learning attitude showed statistically significant differences ($p<0.05$), while other physical fitness indicators did not show statistical significance ($p>0.05$); however, from post-test1 to post-test2 and from baseline to post-test1, all physical fitness and learning attitude variables in the SEM group were statistically significant. In the ST group, between

baseline and post-test2, except for upper body strength, lower body explosiveness, and cardiovascular fitness ($p<0.05$), no other indicators showed significant differences; from baseline to post-test1, except for cardiovascular fitness ($p<0.05$), no other physical fitness and learning attitude variables showed statistical significance ($p>0.05$); from post-test1 to post-test2, except for upper body strength and lower body explosiveness ($p<0.05$), no other physical fitness and learning attitude variables showed statistical significance ($p>0.05$).

Between-group effects showed that there were no statistically significant differences in the dependent variables (physical fitness and learning attitude) at the baseline ($p>0.05$). At post-test1, the differences in cognitive dimension score, affective dimension score, and learning attitude score between the groups were statistically significant ($p<0.05$); at post-test2, the differences in agility, speed, lower body explosiveness, cardiovascular fitness, cognitive dimension score, affective dimension score, and learning attitude score between the groups were statistically significant ($p<0.05$).

These research findings further support that, compared to the ST group, the SEM group showed significant improvements in physical fitness and learning attitude variables (Pritchard et al., 2015; Ward et al., 2017; Wahl-Alexander & Chomentowski, 2018; Nikravan et al., 2019; Fernández-Río et al., 2017; Choi et al., 2021; Zhang et al., 2020).

CHAPTER 5

DISCUSSION, CONCLUSION, IMPLICATION, AND RECOMMENDATION

5.1 Introduction

This chapter discusses the results (physical fitness and learning attitudes), conclusions, implications, and future research recommendations.

5.2 Discussion on Physical Fitness Variables

The first objective of this section is to discuss the results of the seven physical fitness variables in the context of the first objective. The aim is to evaluate the impact of SEM on middle school students' physical fitness and compare and analyse these findings with previous research. Additionally, potential interpretations of the results are considered, along with how these findings can be used to enhance the teaching quality of SEM. The discussion begins with body composition, followed by flexibility, agility, speed, upper body strength, lower body explosiveness, and cardiovascular fitness.

5.2.1 Discussion on Body Composition

The The body composition (BMI-for-age) results indicate that there were no statistically significant differences between the groups at baseline, post-test1, and post-test2 ($p > 0.05$). This finding supports our hypothesis: H01-1: There were no statistically significant differences between the experimental and control groups in body composition at baseline, post-test1, and post-test2 among junior high school students in Hebei, China. Additionally, while the ST

group's BMI-for-age showed no significant changes across the three time points, the SEM group showed no significant difference between baseline and post-test1 but significant differences between baseline and post-test2 and between post-test1 and post-test2 ($p < 0.05$).

These results suggest that SEM has the potential to influence students' BMI-for-age over time, consistent with findings from Segovia & Gutiérrez (2020). However, despite intra-group improvements, no significant between-group differences were observed. While the relatively short intervention duration and moderate exercise intensity may partly explain this, other factors also warrant discussion.

First, student motivation and its variability could have influenced the results. Although SEM is designed to enhance autonomy and engagement, not all students may have responded equally. Some students, particularly those with low intrinsic motivation or a preference for teacher-directed instruction, may have participated passively in SEM sessions, reducing the potential cumulative impact on body composition (Franco et al., 2021).

Second, classroom culture and prior learning habits might have affected how effectively SEM was implemented. In traditional Chinese PE settings, students are accustomed to teacher-led, directive styles. The sudden shift to a more student-centered approach may have caused adjustment difficulties for some students, leading to inconsistent participation quality and less effective physical engagement (Lyu et al., 2023).

Third, differences in extracurricular physical activity and dietary habits could have diluted the observable group-level differences. If both groups engaged in similar levels of physical activity outside of class or maintained stable dietary patterns, the additional impact of SEM would be less pronounced (Liu et al., 2023). Furthermore, growth spurts and puberty-related changes among adolescents can independently influence BMI-for-age, introducing variability unrelated to the intervention (WHO, 2023).

Finally, school-level factors and campus culture may play a role. If the broader school environment does not strongly promote physical activity, even well-designed SEM interventions may be limited in their overall impact (Ward & Griggs, 2011).

In conclusion, while SEM showed intra-group improvements in BMI-for-age, the lack of significant superiority over ST likely results from a combination of intervention duration, student motivation differences, classroom cultural adjustment, external lifestyle factors, and school environmental influences. Future research should integrate motivation-based support strategies, gradually introduce SEM in traditional teaching cultures, and monitor extracurricular habits to better isolate SEM's impact on body composition.

In addition, while this study observed short-term improvements in BMI-for-age within the SEM group, the long-term sustainability of these changes remains uncertain. In real-world educational settings, maintaining positive changes in body composition requires consistent, ongoing engagement and the

integration of health-promoting behaviors beyond the intervention period (Franco et al., 2021). However, sustaining student motivation, ensuring that physical activity becomes a long-term habit, and aligning SEM with school curricula over extended periods present significant challenges (Casey & MacPhail, 2018).

Teachers may face difficulties in continuously implementing SEM due to time constraints, administrative pressures, and the need for specialized pedagogical skills. Moreover, the broader school culture and support from parents and community environments are critical to fostering long-term behavioral change (Perlman & Goc Karp, 2010). Future research should not only extend intervention periods but also explore how to embed SEM into regular school curricula and examine its long-term impact on body composition over multiple academic terms or years. Establishing professional development programs for teachers and creating institutional support mechanisms will be key to achieving sustainable outcomes.

5.2.2 Discussion on Flexibility

The flexibility (sit-and-reach) results show no statistically significant differences between the groups at baseline, post-test1, and post-test2 ($p > 0.05$). This finding supports our hypothesis: H01-2: There were no statistically significant differences between the experimental and control groups in flexibility at baseline, post-test1, and post-test2 among junior high school students in Hebei, China. Additionally, while the ST group showed no significant changes in flexibility across the three time points, the SEM group

showed no significant difference between baseline and post-test1 but significant differences between baseline and post-test2 and between post-test1 and post-test2 ($p < 0.05$).

Although SEM resulted in significant improvements within the group over time, the absence of inter-group differences suggests that factors beyond intervention length need to be considered. First, the specificity of training may have contributed. While SEM sessions are diverse and engaging, they may not have included enough dedicated or progressive flexibility training to surpass the effects of regular physical education (Behm et al., 2016).

Second, student motivation and self-regulation play crucial roles in flexibility development. Some students may have taken the initiative to stretch and practice outside of class due to increased motivation fostered by SEM (Ihle et al., 2021), but this behavior is not guaranteed across all participants. Differences in how students transfer in-class learning to personal practice routines could explain the intra-group improvements without large between-group differences.

Third, classroom culture and peer influence are critical. In a setting where teacher authority traditionally dominates, students may initially lack confidence in leading their own training activities. If classroom culture does not fully support autonomy and peer encouragement, the positive effects of SEM on flexibility development might be limited (Light & Harvey, 2017).

Moreover, external factors such as participation in extracurricular sports or other physical activities may have contributed to flexibility changes in both groups. If students from the control group were involved in other structured or informal physical activities, it would reduce the contrast with the SEM group.

Finally, the school's overall emphasis on physical education and its support for healthy habits may act as a background factor. Without a school-wide culture that values continuous physical self-improvement, even motivated students may face barriers to sustaining stretching routines (Ward & Griggs, 2011).

In summary, the lack of significant between-group differences in flexibility may be due to a combination of factors: insufficient flexibility-specific focus in the intervention, variability in student motivation and self-directed practice, challenges in shifting classroom culture toward autonomy, extracurricular activity dilution, and institutional culture. Future research should integrate flexibility-focused components in SEM, monitor out-of-class practice behaviors, and consider the influence of broader school environments.

Furthermore, although short-term improvements in flexibility were observed within the SEM group, the challenge remains whether these gains can be sustained over time. Long-term flexibility improvements require consistent, focused practice, which may decline once the structured intervention ends (Behm et al., 2016). The sustainability of these outcomes depends on fostering students' intrinsic motivation, encouraging self-directed exercise

habits, and creating a classroom culture that values lifelong physical activity (Ihle et al., 2021).

In practice, maintaining SEM's positive effects on flexibility over semesters or school years may be hindered by scheduling limitations, varying teacher expertise, and insufficient school-level emphasis on holistic fitness (Casey & MacPhail, 2018). Schools must consider implementing supportive policies, integrating SEM principles into ongoing PE curricula, and offering follow-up modules that reinforce flexibility training. Future studies should also investigate how teacher professional development and the cultivation of autonomy-supportive classroom environments influence the long-term effectiveness of SEM in flexibility development.

5.2.3 Discussion on Agility

The agility (arrowhead agility test) results show no statistically significant differences between the groups at baseline and post-test1 ($p > 0.05$), but there is a statistically significant difference at post-test2. This finding refutes our hypothesis: H01-3: There were no statistically significant differences between the experimental and control groups in agility at baseline, post-test1, and post-test2 among junior high school students in Hebei, China. Additionally, while the ST group showed no significant changes in agility across the three-time points, the SEM group showed no significant difference between baseline and post-test1 but significant differences between baseline and post-test2 and between post-test1 and post-test2 ($p < 0.05$). A more detailed explanation follows.

These results indicate that although SEM did not significantly impact agility in the short term (from baseline to post-test1), it had a significant impact over time (by post-test2), suggesting that SEM positively affects students' agility over a more extended period. This finding is consistent with the results of Nikravan et al. (2019). SEM emphasises student engagement and interaction in authentic sports environments, providing structured and contextualised physical education courses. Through diverse exercise and skill training, SEM promotes the comprehensive development of students' physical fitness (Hastie et al., 2011). The model also emphasises teamwork, personal responsibility, and long-term participation, enabling students to improve their reaction speed and agility in constantly changing environments, which is crucial for enhancing agility (Farias et al., 2018).

The differences in the timing of improvements between post-test1 and post-test2 can be explained by the nature of motor learning and physiological adaptation. Parameters such as agility often require more time to develop, as they involve not only muscular strength and speed but also complex neuromuscular coordination, decision-making, and reaction timing (Behm et al., 2010). Early stages of intervention (up to post-test1) may be dominated by cognitive understanding and basic skill familiarization, with minimal observable changes. In contrast, post-test2 outcomes reflect the cumulative effect of continuous practice, where the progression from cognitive and associative learning stages to autonomous execution leads to measurable improvements (Fitts & Posner, 1967). Additionally, physiological adaptations such as neural efficiency and proprioceptive refinement typically occur after

several weeks of repeated exposure (Robinson et al., 2014). This explains why agility showed delayed improvement, while parameters requiring simpler linear strength or flexibility adaptations may respond earlier.

The delayed emergence of significant improvements in agility can also be explained through exercise physiology and neurophysiological adaptation mechanisms. Improvements in agility are not only dependent on muscular strength and speed but also on neuromuscular coordination, proprioceptive response, and the efficiency of the central nervous system in transmitting rapid motor commands (Behm et al., 2010). These adaptations typically require repeated practice and progressive challenges over several weeks. SEM’s continuous and varied practice scenarios likely contributed to motor learning and neuroplastic changes, enabling students to better anticipate, react, and change direction quickly in dynamic game situations (Robinson et al., 2014).

Moreover, from a psychological perspective, SEM’s emphasis on student autonomy and decision-making may enhance cognitive agility, which complements physical agility. The process of making tactical decisions under pressure during team play scenarios trains not only the body but also reaction time and information processing speed, both key components of agility (Light & Harvey, 2017).

Compared to short-term training, long-term, planned physical education is more conducive to improving physical fitness (Miller et al., 2012). Additionally,

the effects of physical training typically manifest after several weeks, especially for complex skills and physical fitness (Robinson et al., 2014). The sustained and structured nature of SEM allows students to undergo progressive physical and cognitive adaptations, which explains why agility improvements were not observed until post-test2.

In contrast, the ST group showed no significant changes in agility throughout the study period, indicating that ST methods have limited effectiveness in improving students' agility. Standard physical education may focus more on basic activities and games, lacking systematic agility training and progressive overload principles, which results in stable agility levels for students (Ward & Griggs, 2011).

Although this study indicates that SEM is effective over a longer period, longer follow-up studies might reveal more comprehensive impacts (Hastie et al., 2011). Future research should consider extending the intervention period to observe the long-term effects of SEM on agility and further explore the neurocognitive and psychological mechanisms that contribute to agility improvements in adolescent students.

While this study demonstrates that SEM can significantly improve agility over a 12-week period, questions remain regarding the long-term sustainability of these improvements. Agility is a complex physical quality that requires continuous practice and progressive challenges to maintain and further develop (Robinson et al., 2014). In real school settings, sustaining SEM

implementation beyond a short intervention period can be complicated by factors such as teacher workload, curriculum rigidity, and competing academic priorities (Light & Harvey, 2017).

Furthermore, as students progress through different grades, changes in academic pressure, school policies, and reduced time for physical education may erode initial gains unless supported by consistent institutional structures and motivational environments. Embedding SEM as part of a continuous school-wide strategy, combined with teacher training, regular performance monitoring, and student feedback mechanisms, will be crucial to achieving lasting agility development (Casey & MacPhail, 2018). Future research should also explore multi-year longitudinal studies to assess the maintenance of agility improvements and identify best practices for long-term SEM integration.

5.2.4 Discussion on Speed

The speed (30-meter sprint test) results show no statistically significant differences between the groups at baseline and post-test1 ($p > 0.05$), but there is a statistically significant difference at post-test2. This finding refutes our hypothesis: H01-4: There were no statistically significant differences between the experimental and control groups in speed at baseline, post-test1, and post-test2 among Junior high school students in Hebei, China. Additionally, while the ST group showed no significant changes in speed across the three-time points, the SEM group showed no significant difference between baseline and post-test1, but significant differences

between baseline and post-test2 and between post-test1 and post-test2 ($p < 0.05$). A more detailed explanation follows.

This result indicates that SEM significantly improves students' speed over extended periods. This finding is similar to previous studies (Wu, 2008; Nikravan et al., 2019). This may be because SEM emphasises physical skills, fitness exercises, and students' overall development and long-term progress (MacPhail et al., 2013). The SEM curriculum provides a more systematic and continuous teaching plan, which helps students show significant improvement over time. In contrast, the ST group showed no significant changes in speed across the three-time points, likely due to the ST methods lacking systematic physical training and long-term planning (Hastie et al., 2011). This is consistent with existing literature, which suggests that ST often significantly struggles to improve students' athletic abilities in the short term (Siedentop et al., 2020).

The observed differences in improvement timelines between post-test1 and post-test2 can be explained by the specific physiological and neurological adaptation processes involved in speed development. Improvements in speed are typically associated with neural adaptations, such as enhanced motor unit recruitment, firing frequency, and synchronization, which require repeated stimulus and sustained training over several weeks (Ross et al., 2001). In the early stages (up to post-test1), students may focus on technique correction and basic posture adjustments, with limited impact on performance. However, as training continues toward post-test2, cumulative

neuromuscular adaptation, muscle fiber recruitment efficiency, and biomechanical refinement become more pronounced, resulting in significant measurable improvements. Additionally, according to the stages of motor learning (Fitts & Posner, 1967), initial improvements are cognitive and slow, while automatic and fluid execution — critical for sprinting performance — emerges after continuous practice and contextual application. This explains why speed improvements become more evident between post-test1 and post-test2 rather than during the initial phase of the intervention.

Although our hypothesis H01-4 showed no significant differences in speed between the experimental and control groups at each time point, the results refuted this hypothesis. A possible reason is that the long-term effects of the SEM curriculum became evident, whereas the short-term intervention did not immediately show differences (MacPhail et al., 2013). Schools should consider adopting the SEM model to continuously improve students' physical fitness and sports participation over extended periods. For educators and policymakers, this study provides strong evidence supporting the application of the SEM curriculum in physical education (Siedentop et al., 2020).

5.2.5 Discussion on Upper Body Strength

The upper body strength (pull-up test) results show no statistically significant differences between the groups at baseline, post-test1, and post-test2 ($p > 0.05$). This finding supports our hypothesis: H01-5: There were no statistically significant differences between the experimental and control groups in upper body strength at baseline, post-test1, and post-test2 among Junior high

school students in Hebei, China. Additionally, while both SEM and ST groups showed no statistically significant differences in upper body strength from baseline to post-test1, both groups showed statistically significant differences from baseline to post-test2 and from post-test1 to post-test2 ($p < 0.05$). A more detailed explanation follows.

These results indicate that although neither the SEM nor ST programs showed significant short-term improvements in students' upper body strength, both groups demonstrated improvements in pull-up test performance over time. This gradual enhancement is likely due to the prolonged participation in physical activities, which gradually strengthens upper body muscles (Faigenbaum et al., 2009). While SEM emphasises skill training and overall development, its short-term effects on specific physical fitness indicators might not be as significant as those achieved through long-term continuous exercise. These results are consistent with existing literature, suggesting that short-term physical interventions may not immediately improve certain specific fitness indicators. Still, long-term sustained exercise can lead to significant improvements (Hoffman et al., 2009). For instance, Faigenbaum et al. (2009) demonstrated that adolescents showed marked improvements in upper body strength after prolonged physical training.

These findings suggest that educators and policymakers should recognise the importance of long-term, sustained physical education programs for developing students' physical fitness, even if short-term results are not immediately apparent. Schools should continue implementing diverse and

systematic physical education programs to improve students' fitness indicators over the long term (Faigenbaum et al., 2009).

5.2.6 Discussion on Lower Body Explosiveness

The lower body explosiveness (vertical jump test) results show no statistically significant differences between the groups at baseline and post-test1 ($p > 0.05$). Still, there is a statistically significant difference in post-test2. This finding refutes our hypothesis: H01-6: There were no statistically significant differences between the experimental and control groups in lower body explosiveness at baseline, post-test1, and post-test2 among Junior high school students in Hebei, China. Additionally, while both SEM and ST groups showed no statistically significant differences in lower body explosiveness from baseline to post-test1, both groups showed statistically significant differences from baseline to post-test2 and from post-test1 to post-test2 ($p < 0.05$). A more detailed explanation follows.

These results indicate that although the SEM program did not significantly improve students' lower body explosiveness between baseline and post-test1, significant progress was observed with more extended intervention. This may be due to the systematic training and repetitive exercises in the SEM curriculum, which enhances lower body explosiveness over time (Harrison & Gaffney, 2001). The SEM program emphasises long-term development and continuous training, which may take longer to show significant effects on specific physical fitness indicators (Hastie et al., 2011). In contrast, the ST group's progress was relatively smaller, likely due

to the ST methods' lack of systematisation and continuity (Siedentop et al., 2020). These results are consistent with existing literature, indicating that SEM significantly improves students' physical fitness with long-term intervention (Harrison & Gaffney, 2001). For example, Hastie et al. (2011) found that systematic physical education programs can significantly improve students' athletic performance and physical fitness.

Furthermore, the intervention frequency in this study was three times per week, consistent with some previous studies. For instance, Faigenbaum et al. (2009) also used a frequency of three times per week, significantly improving adolescents' physical fitness after long-term intervention. Additionally, Harrison and Gaffney (2001) found that sustained training three times per week effectively improved students' physical fitness. However, other studies with different intervention frequencies have shown varying effects. For example, Hoffman et al. (2009) found that although an intervention frequency of twice per week had some short-term impact on strength and explosiveness, the long-term effects were less significant than those of three or more times per week. Conversely, a frequency of five times per week may lead to overtraining, resulting in less significant effects than three times per week, as students might not fully recover, risking fatigue and injury (Hoffman et al., 2009).

This study provides strong evidence for educators and policymakers to support the application of the SEM program in physical education (Siedentop

et al., 2020). Future research could explore different intervention frequencies to compare their effects on physical fitness improvement.

5.2.7 Discussion on Cardiovascular Fitness

The cardiovascular fitness (YO-YO test) results show no statistically significant differences between the groups at baseline and post-test1 ($p > 0.05$). Still, there is a statistically significant difference in post-test2. This finding refutes our hypothesis: H01-7: There were no statistically significant differences between the experimental and control groups in cardiovascular fitness at baseline, post-test1, and post-test2 among Junior high school students in Hebei, China. Additionally, although both SEM and ST groups showed no statistically significant differences in cardiovascular fitness from baseline to post-test1 and from post-test1 to post-test2, both groups showed statistically significant differences from baseline to post-test2 ($p < 0.05$). A more detailed explanation follows.

These results indicate that although the SEM program did not significantly improve students' cardiovascular fitness between baseline and post-test1, significant progress was observed with more extended intervention. This aligns with previous research findings (Pritchard et al., 2015; Ward et al., 2017; Wahl-Alexander & Chomentowski, 2018; Wu, 2008; Nikravan et al., 2019; Menickelli & Hastie, 2014). This improvement may be due to the purposeful and sustained practice in the SEM curriculum, which enhances cardiovascular function over time (Ward et al., 2017). Another possible explanation is that the intrinsic structure of SEM creates a highly autonomous

environment in physical education (Wahl-Alexander & Chomentowski, 2018). Compared to direct instruction, these key features—including team relationships, festive atmospheres, and peer competition—provide motivated students with greater enjoyment (Wallhead et al., 2014). These results support the hypothesis that physical education moves many students with varying levels of motivation along a self-determined continuum (Hastie & Wallhead, 2016). This perspective is also reflected in the ST group in this study, whose smaller improvement may be due to the ST method's rigid control over content, giving students a sense of being controlled, which can negatively impact their physical fitness development (Xie, 2012).

These findings suggest schools should consider adopting the SEM model to enhance students' cardiovascular fitness over extended periods. For educators and policymakers, this study provides strong evidence supporting the application of SEM programs in physical education (Siedentop et al., 2020). However, some scholars suggest that given the potential benefits demonstrated in this context, it is necessary to replicate the research in different settings with larger sample sizes to evaluate SEM's broader applicability and effectiveness (Wahl-Alexander & Chomentowski, 2018).

5.3 Discussion on Learning Attitude

The main objective of this section is to discuss the two variables included in learning attitude (cognitive learning attitude and affective learning attitude) and the overall results of learning attitude against the backdrop of the second objective. The purpose is to evaluate the impact of SEM on Junior high school

students' learning attitudes and compare and analyse previous research results. Additionally, considerations are given to possible explanations for the results and how to utilise these findings to enhance the teaching quality of SEM.

5.3.1 Discussion on Cognitive Learning Attitude Dimension Score

The cognitive learning attitude, measured by the cognitive dimension score, demonstrated statistically significant differences between the baseline, post-test1, and post-test2 groups. This finding contradicts our research hypothesis (H02-1), positing no statistically significant differences between the experimental and control groups in cognitive learning attitude at these time points among Junior high school students in Hebei, China. Furthermore, students in the SEM group exhibited statistically significant within-group differences in cognitive dimension scores across all three time periods (baseline to post-test1 and post-test1 to post-test2), while those in the ST group did not show significant differences across the same periods.

SEM's emphasis on providing structured and contextualised sports programs may effectively promote students' cognitive learning attitudes. By stressing teamwork, individual responsibility, and long-term engagement, SEM facilitates the development of cognitive abilities and learning attitudes in real sports contexts. Simulating real games and scenarios enables students to enhance their cognitive abilities and interest in learning, which are crucial for improving cognitive learning attitudes. Additionally, SEM's teaching methods may increase students' learning engagement and motivation, as students are

more likely to invest more tremendous effort in highly engaging environments, leading to more significant progress in cognition and attitudes.

In contrast, the ST group showed no significant change in cognitive learning attitudes throughout the study period, indicating that standard teaching methods have limited effectiveness in enhancing students' cognitive learning attitudes. Standard physical education may prioritise basic activities and games, but it needs more systematic cultivation of mental abilities and learning attitudes. This approach may fail to stimulate students' interest and enthusiasm for learning, resulting in stable cognitive learning attitudes.

Therefore, SEM significantly improves students' cognitive learning attitudes. However, future research should extend the duration of the SEM intervention or increase follow-ups after the experiment to observe its long-term effects. Additionally, increasing the sample size would enhance the generalizability of the results. Studies with larger sample sizes can provide more representative conclusions, improving research results' reliability and external validity.

5.3.2 Discussion on Affective Learning Attitude Dimension Score

The affective learning attitude, as measured by the affective dimension score, demonstrated statistically significant differences between the baseline and post-test1 and between post-test1 and post-test2 groups. This finding contradicts our research hypothesis (H02-2), which posited no statistically significant differences between the experimental and control groups in affective learning attitudes at these time points among Junior high school

students in Hebei, China. Furthermore, students in the SEM group exhibited statistically significant within-group differences in affective dimension scores across all three time periods (baseline to post-test1 and post-test1 to post-test2), while those in the ST group did not show significant differences across the same periods.

SEM effectively promotes students' affective learning attitudes, consistent with previous research findings. SEM's teaching philosophy emphasises the team learning process, enhancing student interaction and a sense of belonging in experimental classes. This approach deepens students' understanding of teamwork, particularly encouraging those reluctant to participate or disengaged from classroom activities to join in physical activities. This further enlivens the classroom atmosphere and enhances the enjoyment of physical education. Creating a pleasant atmosphere in physical education classes is essential, as enjoyment of physical education correlates positively with enjoyment of physical activities.

In contrast, the ST group showed no significant change in affective learning attitudes throughout the study period, indicating that standard teaching methods have limited effectiveness in improving students' affective learning attitudes. Standard physical education may prioritise basic activities and games, needing more systematic cultivation of affective abilities and learning attitudes. This approach may fail to stimulate students' emotional engagement, resulting in stable, effective learning attitudes.

However, most existing studies have yet to consider other factors influencing students' extracurricular activities and family economic status, which may affect the results. Future research could control for these variables to assess the impact of SEM on affective learning attitudes more accurately.

5.3.3 Discussion on Holistic Learning Attitude Score

The learning attitude, as indicated by the learning attitude score, demonstrated statistically significant differences between the post-test1 and post-test2 groups, contrary to our research hypothesis (H02-3) of no significant differences between the experimental and control groups at baseline, post-test1, and post-test2. Additionally, students in the SEM group exhibited statistically significant within-group differences in learning attitude scores across all three periods (baseline to post-test1 and post-test1 to post-test2), while those in the ST group did not show significant differences across the same periods.

SEM effectively promotes students' learning attitudes, consistent with previous research findings. Participating in SEM projects cultivates a more positive learning attitude in students. SEM provides students with equal opportunities to play and practice with teammates during the long season, fostering close relationships among students and experiencing the joy of games, stimulating interest in learning and leading to better learning attitudes. SEM's teaching methods may also increase student engagement and motivation. Students are more likely to invest more emotionally in environments with higher levels of participation, leading to greater progress in

attitudes and motivation. SEM's fun and challenging activities inspire students' interest and enthusiasm for learning, significantly improving their learning attitudes relatively quickly.

In contrast, the ST group showed no significant change in learning attitude throughout the study period, with some studies even suggesting a deterioration in learning attitude. This difference may be attributed to SEM providing students with opportunities to make friends and experience the joy of games, enhancing motivation for learning. Intimacy among team members is critical to strengthening learning motivation, and students' role-playing within teams, such as captains, coaches, scorers, referees, and players, allows them to develop a sense of responsibility. Pursuing higher scores for the team promotes trust and teamwork among members. In SEM groups, most high-performing students are willing to help lower-performing teammates, and some even conduct further research on the content they have learned. Due to team pride, role responsibilities, and teammates' gratitude, lower-performing students strive to improve themselves. Students experience the fun of games, sports, and friendship through role-playing. A relaxed and friendly learning atmosphere fosters a positive learning attitude in football. In contrast, students in the ST group passively receive knowledge and skills from teachers, leading to a less positive attitude toward physical education classes.

Therefore, SEM significantly improves students' learning attitudes. Future research should focus on sample selection in intervention plans, as most

studies are conducted in the same school. Additionally, the theoretical basis of surveys on learning attitudes among junior high school students could be stronger, with few researchers using questionnaires based on attitude theory (primarily cognitive and affective dimensions) to study students' learning attitudes.

5.4 Conclusion

The results demonstrate that the SEM group significantly enhanced Junior high school students' physical fitness and learning attitudes. Moreover, the SEM group outperformed the ST group significantly in physical fitness indicators such as agility, speed, lower body explosiveness, and cardiovascular fitness, as well as in cognitive learning attitude, affective learning attitude, and overall learning attitude score. Therefore, our hypotheses H01-3, H01-4, H01-6, H01-7, H02-1, H02-2, and H02-3 were rejected, while H01-1, H01-2, and H01-5 were accepted.

5.4.1 Physical Fitness

The intervention employed by SEM, consisting of three sessions per week, lasting 40 minutes, demonstrated significant effects on improving students' BMI-for-age, flexibility, agility, speed, upper body strength, lower body explosiveness, and cardiovascular fitness throughout 6 to 12 weeks. In contrast, the ST group showed significant improvements only in upper body strength and lower body explosiveness during the 6 to 12-week period, along with cardiovascular fitness in the initial six weeks. These findings suggest that the SEM intervention, conducted three times a week with 40-minute sessions,

serves as an effective instructional model for enhancing physical fitness among Junior high school students and outperforms ST. However, it is advisable to extend the duration of SEM intervention to observe its long-term effects. A more comprehensive assessment of SEM's impact on student physical fitness can be conducted through prolonged intervention. Furthermore, replicating the study in different contexts and with larger sample sizes is warranted.

5.4.2 Learning Attitude

The intervention employed by SEM, consisting of three sessions per week with each session lasting 40 minutes, demonstrated significant effects on improving students' cognitive learning attitude, affective learning attitude, and overall learning attitude across three time periods: Baseline - post-test1, Baseline - post-test2, and post-test1 - post-test2. In contrast, ST showed no significant improvement in these aspects. These findings suggest that the SEM intervention, conducted three times a week with 40-minute sessions, serves as an effective instructional model for enhancing learning attitudes among Junior high school students and outperforms ST. SEM emphasises structured and contextualised physical education courses, emphasising teamwork, personal responsibility, and long-term engagement. By simulating real games and scenarios, students can enhance their cognitive abilities and interest in learning in constantly changing environments, crucial for improving student learning attitudes.

5.5 Research Implication

This study contributes to the growing body of research on the SEM by providing evidence of its effectiveness in improving certain aspects of students' physical fitness and learning attitudes in the short term. However, the implications extend beyond these immediate findings and highlight several areas that require attention for future research and practical application.

Firstly, there is a clear need for longitudinal studies to assess the long-term impact and sustainability of SEM interventions. While this study focused on a 12-week intervention period, behavioral and physiological changes — particularly those related to body composition, flexibility, and agility — may only manifest fully over extended periods. Longitudinal research would allow scholars and educators to explore how SEM influences students' physical development, motivation, and learning outcomes across multiple academic years, and whether these benefits are maintained or decline without continued structured implementation. Such studies should also investigate the factors that contribute to sustained engagement and identify strategies for embedding SEM into school culture over the long term.

Secondly, the scalability of SEM across different educational systems, regions, and demographic groups warrants further exploration. While this study was conducted among junior high school students in Hebei Province, China, its applicability in other provinces with varying socioeconomic

conditions, educational policies, and school resources remains uncertain. Moreover, the adaptability of SEM in different international educational contexts, where classroom cultures and pedagogical traditions may differ significantly, should be carefully examined. Researchers and policymakers need to consider how local educational structures, teacher training standards, and community attitudes toward physical education influence the successful adoption of SEM.

Finally, from a policy and practice perspective, this study highlights the importance of developing support systems — including teacher professional development, curriculum integration frameworks, and resource allocation — to ensure SEM can be implemented at scale and maintained over time. School administrators and education authorities should consider multi-level strategies to facilitate SEM integration, particularly in under-resourced or rural areas where logistical and cultural barriers may be more pronounced.

In conclusion, future research should not only expand SEM applications through longitudinal, multi-regional, and cross-cultural studies but also address the institutional and systemic factors that affect its scalability and sustainability. Only by understanding these complexities can SEM realize its full potential in improving physical education outcomes globally.

5.6 Recommendations for Future Studies

Based on the findings of this study, several recommendations for future research are proposed in terms of population, intervention, and outcomes:

(a) Population: Since this study included a mixed-gender sample of junior high school students, future research could examine the effects of the SEM on students of different genders. Due to physiological and psychological differences between male and female adolescents, such as variations in physical fitness levels and learning attitudes, SEM interventions may yield different results across genders. Conducting gender-specific studies would help clarify these potential differences, enabling researchers and educators to tailor SEM interventions more effectively. By customising SEM approaches based on gender-specific needs, future studies could enhance the model's impact on both physical fitness and learning attitudes, thereby maximising its educational value for students.

(b) Intervention: Recent systematic reviews have identified the emergence of hybrid instructional models that incorporate SEM as a foundational framework, such as the SEM-TGfU (Teaching Games for Understanding) model. This combination leverages SEM’s emphasis on authentic sports experiences while integrating TGfU's focus on tactical understanding and decision-making skills. However, research on these hybrid models is still in its early stages, with many areas yet to be explored. Future studies could investigate the effects of such hybrid models on students' learning attitudes and physical fitness. By utilising SEM as a core structure while incorporating complementary approaches like TGfU, researchers could provide students with a more holistic and engaging learning experience. This blended approach may offer broader benefits by addressing both physical

development and cognitive engagement, ultimately enhancing students' overall experience in physical education.

(c) Outcomes: This study did not investigate the potential interaction between learning attitudes and physical fitness. Future research should consider exploring this relationship, as learning attitudes and physical fitness may reinforce each other. For example, a positive attitude toward PE could enhance students' engagement and persistence, potentially leading to improved physical fitness outcomes. Conversely, enhanced physical fitness might boost students' confidence and satisfaction in PE classes, fostering a more positive learning attitude. Investigating this interplay could provide a deeper understanding of how the SEM impacts students holistically, revealing whether interventions that target both domains simultaneously might yield synergistic benefits.

Furthermore, beyond the improvements in learning attitudes and physical fitness observed in this study, it was also noted that students' motivation to participate in PE increased. Future studies should also consider measuring additional psychological variables, such as motivation, self-efficacy, and enjoyment to enhance the generalisability of SEM research findings. Including these psychological factors could provide a more comprehensive view of SEM's impact on student outcomes. By examining how SEM influences both physical and psychological dimensions, researchers can better assess its effectiveness as an educational model, contributing valuable insights for its application in academic settings. Such findings could

strengthen the evidence base for SEM, demonstrating its potential to improve physical and cognitive outcomes and foster positive psychological growth, thereby supporting students' overall development in PE programs.

REFERENCES

Abdullaeva, B. P. (2021). Football as a means of physical education. *Current Research Journal of Pedagogics, 2*(08), 72-76.

Aboshkair, K. A. A. (2012). Effect of the implementation levels of the physical education program on health-related physical fitness of children in Selangor, Malaysia (Doctoral dissertation, Universiti Putra Malaysia).

Aboshkair, K. A. A., Saidon, B. A., Kok, L. Y., & Bahaman, B. A. S. (2012). Factors affecting levels of health-related physical fitness in secondary school students in Selangor, Malaysia. *Journal of Basic & Applied Sciences, 8*, 202–216.

Ajzen, I., Fishbein, M., Lohmann, S., & Albarracín, D. (2018). The influence of attitudes on behavior. In D. Albarracín & B. T. Johnson (Eds.), *The handbook of attitudes, volume 1: Basic principles* (pp. 197-255). Routledge.

Albaladejo-Saura, M., Vaquero-Cristóbal, R., García-Roca, J. A., & Esparza-Ros, F. (2022). Influence of biological maturation status on selected anthropometric and physical fitness variables in adolescent male volleyball players. *PeerJ, 10*, e13216. https://doi.org/10.7717/peerj.13216

Alkhawaldeh, A., AlBashtawy, M., Rayan, A., Abdalrahim, A., Musa, A., Eshah, N., & Khraisat, O. (2023). Application and use of Andersen's behavioral model as theoretical framework: A systematic literature review from 2012–2021. *Iranian Journal of Public Health, 52*(7), 1346–1358.

Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *Handbook of social psychology* (pp. 798-844). Clark University Press.

Anderson, C., & Sally, D. (2013). *The numbers game: Why everything you know about soccer is wrong*. Penguin Books. https://doi.org/10.33448/rsd-v10i8.17824

Bagozzi, R. P., & Burnkrant, R. E. (1979). Attitude organization and the attitude–behavior relationship. *Journal of Personality and Social Psychology, 37*(6), 913–929. https://doi.org/10.1037/0022-3514.37.6.913

Bai, X., Xiao, W., Soh, K. G., Agudamu, & Zhang, Y. (2023). 12-week concurrent brisk walking and Taijiquan (Tai Chi) improve balance, flexibility, and muscular strength of Chinese older women. *PLOS ONE*, 18(10), e0293483. https://doi.org/10.1371/journal.pone.0293483

Bangsbo, J., Iaia, F. M., & Krustrup, P. (2008). The Yo-Yo Intermittent Recovery Test: A useful tool for evaluation of physical performance in

intermittent sports. *Sports Medicine, 38*(1), 37–51. https://doi.org/10.2165/00007256-200838010-00004

Baquet, G., et al. (2001). Flexibility in youth: A systematic review. *European Journal of Applied Physiology*.

Barba-Martín, R. A., Bores-García, D., Hortigüela-Alcalá, D., & González-Calvo, G. (2020). The application of the teaching games for understanding in physical education: Systematic review of the last six years. *International Journal of Environmental Research and Public Health*, 17(9), 3330.

Barker, D., McElduff, P., D'Este, C., & Campbell, M. J. (2016). Stepped wedge cluster randomized trials: A review of the statistical methodology used and available. *BMC Medical Research Methodology*, 16(1), 1-19. https://doi.org/10.1186/s12874-016-0176-5

Beni, S., Chróinín, D. N., & Fletcher, T. (2021). 'It's how PE should be!': Classroom teachers' experiences of implementing Meaningful Physical Education. *European Physical Education Review, 27*(3), 666-683. https://doi.org/10.1177/1356336X209841

Bessa, C., Hastie, P., Araújo, R., & Mesquita, I. (2019). What do we know about the development of personal and social skills within the Sport Education Model: A systematic review? *Journal of Sports Science and Medicine, 18*(4), 812-829.

Bessa, C., Hastie, P., Rosado, A., & Mesquita, I. (2021). Sport education and traditional teaching: Influence on students' empowerment and self-confidence in high school physical education classes. *Sustainability, 13*(2), 578. https://doi.org/10.3390/su13020578

Bolanowski, M., & Nilsson, B. E. (2001). Assessment of human body composition using dual-energy x-ray absorptiometry and bioelectrical impedance analysis. *Medical Science Monitor*, 7(5), 1029-1033.

Bores-García, D., Hortigüela-Alcalá, D., Fernandez-Rio, F. J., González-Calvo, G., & Barba-Martín, R. (2021). Research on cooperative learning in physical education: Systematic review of the last five years. *Research Quarterly for Exercise and Sport*, 92(1), 146-155.

Borga, M., West, J., Bell, J. D., Harvey, N. C., Romu, T., Heymsfield, S. B., & Dahlqvist Leinhard, O. (2018). Advanced body composition assessment: From body mass index to body composition profiling. *Journal of Investigative Medicine*, 66(5), 1-9.

Bouchard, C., Antunes-Correa, L. M., Ashley, E. A., Franklin, N., Hwang, P. M., Mattsson, C. M., & Wheeler, M. T. (2015). Personalized preventive medicine: Genetics and the response to regular exercise in preventive

interventions. *Progress in Cardiovascular Diseases*, 57(4), 337-346. https://doi.org/10.1016/j.pcad.2014.10.004

Bowler, M., Sammon, P., & Casey, A. (2022). Health-based physical education: A pedagogical model in focus. In M. Bowler, P. Sammon, & A. Casey (Eds.), *Physical education pedagogies for health* (pp. 62-76). Routledge.

Burhaein, E., Ibrahim, B. K., & Pavlovic, R. (2020). The relationship of limb muscle power, balance, and coordination with instep shooting ability: A correlation study in under-18 football athletes. *International Journal of Human Movement and Sports Sciences*, 8(5), 265-270.

Cai, M. (2023). *Chengdu sixth-grade primary school students' physical activity levels, the relationship between physical self-esteem and anxiety* (Master's dissertation, Chengdu University). https://doi.org/10.27917/d.cnki.gcxdy.2023.000207

Caldbeck, P. M. (2020). Contextual sprinting in Premier League football (Doctoral thesis, Liverpool John Moores University).

Cao, K. P., Yin, X. Q., Yang, G. Q., & Li, B. G. (2018). Regional applicability of the Chinese version of the Youth Physical Activity Questionnaire (PAQ-A). *Contemporary Sports Technology, 2*, 241–243. https://doi.org/10.16655/j.carolcarroll.nki.2095-2813.2018.02.241

Carriedo, A., Méndez Giménez, A., Fernández Río, J., & Cecchini, J. (2020). Teaching and assessment of physical fitness in secondary education: A proposal based on the sport education model. *Cultura_Ciencia_Deporte, 15*(46), 31-37. https://doi.org/10.12800/ccd.v15i46.1534

Casey, A., Goodyear, V. A., & Dyson, B. (2016). Models-based practice in physical education: Theory, research, and practice. Physical Education and Sport Pedagogy, 21(3), 285-298. https://doi.org/10.1080/17408989.2015.1072507

Castillo, D., Raya-González, J., Sarmento, H., Clemente, F. M., & Yanci, J. (2021). Effects of including endurance and speed sessions within small-sided soccer games periodization on physical fitness. *Biology of Sport*, 38(2), 291.

Chalil, D. H., Febrianty, M. F., & Sartono, H. (2018). The validity and reliability of Arrowhead Agility Test in football. *Journal of Physical Education and Sport, 18*(1), 181–187.

Chen, B. (2017). A comparative study of the "Sport Education" model and the "Three Basics" physical education teaching model. *Sports Technology*, (04), 130-131+133. https://doi.org/10.14038/j.cnki.tykj.2017.04.097

Chen, B. (2017). Comparative study of "Sport Education" mode and "Three basic" PE teaching mode. Sports Science and Technology, 38(4), 130-131 + 133.

Chen, S., & Gu, X. (2018). Effects of cardiorespiratory fitness and weight status on knowledge of physical activity and fitness, attitude toward physical education, and physical activity. *BMC Public Health, 18*, 1-7. https://doi.org/10.1186/s12889-018-5372-5

Chiel, H. J., & Beer, R. D. (1997). The brain has a body: Adaptive behavior emerges from interactions of nervous system, body, and environment. *Trends in Neurosciences, 20*(12), 553-557. https://doi.org/10.1016/S0166-2236(97)01150-1

Chu, Y., Chen, C., Wang, G., & Su, F. (2022). The effect of education model in physical education on student learning behavior. Frontiers in psychology, 13, 944507. https://doi.org/10.3389/fpsyg.2022.944507

Cooper, B., Eva, N., Fazlelahi, F. Z., Newman, A., Lee, A., & Obschonka, M. (2020). Addressing common method variance and endogeneity in vocational behavior research: A review of the literature and suggestions for future research. *Journal of Vocational Behavior*, 121, 103472.

Corbin, C. B., & Lindsey, R. (1997). *Concepts of physical fitness* (12th ed.). McGraw-Hill.

Cuevas, R., García-López, L. M., & Serra-Olivares, J. (2016). Sport education model and self-determination theory: An intervention in secondary school children. *Kinesiology, 48*(1), 30-38. https://doi.org/10.26582/k.48.1.15

Dalen, T., Sandmæl, S., Stevens, T. G., Hjelde, G. H., Kjøsnes, T. N., & Wisløff, U. (2021). Differences in acceleration and high-intensity activities between small-sided games and peak periods of official matches in elite soccer players. *The Journal of Strength & Conditioning Research*, 35(7), 2018-2024.

de Almeida Andrade, R. (2020). Physical fitness and cardiovascular health in college students: A cross-sectional study (Doctoral dissertation, Universidade de Trás-os-Montes e Alto Douro).

Delgado-Floody, P., Cristi-Montero, C., Jerez-Mayorga, D., Ruiz-Ariza, A., Guzmán-Guzmán, I. P., Álvarez, C., ... & Caamaño-Navarrete, F. (2024). Exploring the mediating role of promoting school physical activity on the relationship between low socioeconomic status and academic achievement and school climate: Evidence from 4,990 Chilean schools. Frontiers in Public Health, 12, 1426108. https://doi.org/10.3389/fpubh.2023.1426108

Demers, D. (2015). Using the FITT principle to examine the relationship between exercise and sexual performance indicators. *International Journal of Health Sciences, 3*(4), 79-93.

Dong, Y., Lau, P. W., Dong, B., Zou, Z., Yang, Y., Wen, B., ... & Patton, G. C. (2019). Trends in physical fitness, growth, and nutritional status of Chinese children and adolescents: A retrospective analysis of 1.5 million students from six successive national surveys between 1985 and 2014. *The Lancet Child & Adolescent Health, 3*(12), 871-880. https://doi.org/10.1016/S2352-4642(19)30302-5

Donner, A., Birkett, N., & Buck, C. (1981). Randomization by cluster: Sample size requirements and analysis. *American Journal of Epidemiology, 116*(4), 906–914. https://doi.org/10.1093/oxfordjournals.aje.a113261

Duan, J., Hu, H., Wang, G., & Arao, T. (2015). Study on current levels of physical activity and sedentary behavior among middle school students in Beijing, China. *PLoS ONE*, 10(7), e0133544.

Elamin, A. (2005). Assessment of nutritional status, College of Medicine. Sultan Qaboos University, Muscat, Oman.

Faigenbaum, A. D., Kraemer, W. J., Blimkie, C. J., Jeffreys, I., Micheli, L. J., Nitka, M., & Rowland, T. W. (2009). Youth resistance training: Updated position statement paper from the national strength and conditioning association. *Journal of Strength and Conditioning Research, 23*(Suppl 5), S60-S79. https://doi.org/10.1519/JSC.0b013e31819df407

Fang, L. (2024). An Empirical Study on the Effect of Sport Education Model on enhancing the core sports literacy of junior high school students (Master's Thesis). Qufu Normal University. https://doi.org/10.27267/d.cnki.gqfsu.2024.001268

Farias, C. F., Hastie, P. A., & Mesquita, I. (2018). Scaffolding student-coaches' instructional leadership toward student-centred peer interactions: A yearlong action-research intervention in sport education. *European Physical Education Review, 24*(3), 290-307. https://doi.org/10.1177/1356336X16684108

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fazio, R. H., & Zanna, M. P. (1981). Direct experience and attitude–behavior consistency. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 161–202). Academic Press.

Fernandez-Rio, J., & Iglesias, D. (2024). What do we know about pedagogical models in physical education so far? An umbrella review.

Physical Education and Sport Pedagogy, 29(2), 190-205. https://doi.org/10.1080/17408989.2023.1234567

Fernández-Río, J., Méndez-Giménez, A., & Méndez Alonso, D. (2017). Effects of two instructional approaches, sport education and direct instruction, on secondary education students' psychological responses. https://doi.org/10.6018/300561

Franco, E., Tovar, C., González-Peño, A., & Coterón, J. (2021). Effects of a sport education model-based teaching intervention on students' behavioral and motivational outcomes within the physical education setting in the COVID-19 scenario. *Sustainability (Switzerland)*, 13(22).

Gao H. (2020). Comparative study of Sport Education Mode and traditional Physical Education Mode. *Physical Education*, 40(01), 40-42.

Garramona, F. T., Moreira, R. S. O., Korch, H., Cruz, J. E., & Teixeira, L. F. M. (2024). Enhancing flexibility through Chinese auriculotherapy: Investigating the impact on sit and reach test–A randomized controlled trial. *Traditional and Integrative Medicine*.

Giménez-Meseguer, J., Ferriz-Valero, A., & Baena-Morales, S. (2022). Impact of sport education model on sports lifestyle and attitudes of vocational education training students. *Education Sciences, 12*(12), 896. https://doi.org/10.3390/educsci12120896

Gokhale, R., & Kirschner, B. S. (2003). Assessment of growth and nutrition. Best Practice & Research Clinical Gastroenterology, 17(2), 153-162. https://doi.org/10.1016/S1521-6918(02)00103-2

Gonzàlez, V. E. (1992). *On human attitudes: Root metaphors in theoretical conceptions*. Göteborg, Sweden: Vasastadens Bokbinderi AB.

Grove, S. K. (2018). Understanding statistics in research. *Understanding Nursing Research E-Book: Building an Evidence-Based Practice*, 292.

Gubacs-Collins, K. (2015). The socratic gymnasium: Learning lessons of life through physical education. *The Physical Educator*, 72, 76-98.

Guijarro, E., Rocamora, I., González-Víllora, S., & Arias-Palencia, N. M. (2020). The role of physical education in the achievement of international recommendations: A study based on pedagogical models. *Journal of Human Sport and Exercise*, 15(4), 849-860.

Hajian Tilaki, K. O., Sojjodi, P., & Razavi, A. (2011). Prevalence of overweight and obesity and associated risk factors in urban primary-school children in Babol, Islamic Republic of Iran. EMHJ-Eastern Mediterranean Health Journal, 17(2), 109-114.

Harrison, J. M., & Gaffney, A. B. (2001). *Differentiated instruction in physical education*. Human Kinetics.

Harvey, S., & Jarrett, K. (2014). A review of the game-centred approaches to teaching and coaching literature since 2006. *Physical Education and Sport Pedagogy, 19*(3), 278-300. https://doi.org/10.1080/17408989.2012.754005

Hastie, P. A., & Casey, A. (2014). Fidelity in models-based practice research in sport pedagogy: A guide for future investigations. Journal of Teaching in Physical Education, 33(3), 422-431. https://doi.org/10.1123/jtpe.2013-0133

Hastie, P. A., & Trost, S. G. (2002). Student physical activity levels during a season of sport education. Pediatric Exercise Science, 14(1), 64-74. https://doi.org/10.1123/pes.14.1.64

Hastie, P. A., & Wallhead, T. (2016). Models-based practice in physical education: The case for sport education. *Journal of Teaching in Physical Education, 35*(4), 390-399. https://doi.org/10.1123/jtpe.2016-0092

Hastie, P. A., Buchanan, A. M., Wadsworth, D. D., & Sluder, B. J. (2009). The impact of an obstacle course sport education season on students' aerobic fitness levels. *Research Quarterly for Exercise and Sport, 80*(4), 788-791.

Hastie, P. A., Calderón, A., Rolim, R. J., & Guarino, A. J. (2013). The development of skill and knowledge during a sport education season of track and field athletics. *Research Quarterly for Exercise and Sport*, 84(3), 336-344. https://doi.org/10.1080/02701367.2013.813894

Hastie, P. A., Martinez, D., & Calderón, A. (2011). A review of research on sport education: 2004 to the present. *Physical Education and Sport Pedagogy, 16*, 103–132. https://doi.org/10.1080/17408989.2010.535202

Havas, E. (2021). Experiences of the novice instructors of the Teaching Personal and Social Responsibility (TPSR) training program, model and transfer of learning (Master's thesis, University of Jyväskylä).

He, J. (2023). Experimental research on Sport Education Model in junior high school football teaching (Master's Thesis). Shenyang Sports College. https://doi.org/10.27329/d.cnki.gstyc.2023.000067

He, S. (2021). Effects of nature exposure intervention on self-regulation and decision-making in mentally-fatigued university soccer players (Doctoral thesis). Universiti Putra Malaysia.

Hilgard, E. R. (1980). The trilogy of mind: Cognition, affection, and conation. *Journal of the History of Behavioral Sciences, 16*(2), 107-117. https://doi.org/10.1002/1520-6696(198004)16:2<107:AID-JHBS2300160204>3.0.CO;2-B

Hoffman, J. R., Ratamess, N. A., Faigenbaum, A. D., & Kang, J. (2009). Effects of a short-term resistance training program on strength and power in adolescents. *Journal of Strength and Conditioning Research, 23*(3), 823-831. https://doi.org/10.1519/JSC.0b013e3181a02a77

Hu, H., Duan, J., Wang, G., & Arao, T. (2014). Reliability and validity of a Chinese version of the students' attitudes toward physical education scale and its related factors. *Advances in Physical Education*, 4(4), 181.

Hu, H., Duan, J., Wang, G., & Arao, T. (2014). Reliability and validity of a Chinese version of the students' attitudes toward physical education scale and its related factors. *Advances in Physical Education, 4*(4), 181–189. https://doi.org/10.4236/ape.2014.44022

Huang, S., Fiero, M. H., & Bell, M. L. (2016). Generalized estimating equations in cluster randomized trials with a small number of clusters: Review of practice and simulation study. *Clinical Trials*, 13(4), 445-449. https://doi.org/10.1177/1740774516643498

Hutchins, E. (1996). *Cognition in the wild* (2nd ed.). MIT Press.

Iglesias, D., Fernandez-Rio, J., & Rodriguez-Gonzalez, P. (2022). Moderate-to-vigorous physical activity in physical education: A review of reviews. *Journal of Teaching in Physical Education, 42*(4), 640-646.

Insko, C. A., & Schopler, T. (1972). *Experimental social psychology*. Academic Press.

Iserbyt, P., Dehandschutter, T., Leysen, H., & van der Mars, H. (2022). The effect of a coaching clinic on student-coaches' behavior during a basketball Sport Education season with preservice teachers. Journal of Teaching in Physical Education, 42(3), 563-572. https://doi.org/10.13189/saj.2023.100219.

Jakicic, J. M. (2019). Physical activity: A key lifestyle behavior for prevention of weight gain and obesity. *American College of Sports Medicine*. https://www.acsm.org/blog-detail/acsm-certified-blog/2019/07/22/physical-activity-prevention-weight-gain-obesity

Jeejeebhoy, K. N. (1998). Nutritional assessment. Gastroenterology Clinics of North America, 27(2), 347-369. https://doi.org/10.1016/S0889-8553(05)70338-2

Kail, R. V., & Cavanaugh, J. C. (2007). Human development: A life-span view (4th ed.). Australia: Thomson Wadsworth.

Kang, Y., & Gao, R. (2016). A comparison of teaching objectives between the Sport Education model and the traditional physical education teaching model. *Sports*, (05), 70-71.

Karlsen, T., Aamot, I. L., Haykowsky, M., & Rognmo, Ø. (2017). High-intensity interval training for maximizing health outcomes. *Progress in Cardiovascular Diseases, 60*(1), 67-77. https://doi.org/10.1016/j.pcad.2017.03.006

Keiner, M., Kadlubowski, B., Sander, A., Hartmann, H., & Wirth, K. (2022). Effects of 10 months of speed, functional, and traditional strength training on strength, linear sprint, change of direction, and jump performance in trained adolescent soccer players. *The Journal of Strength & Conditioning Research, 36*(8), 2236–2246. https://doi.org/10.1519/JSC.0000000000003807

Krech, D., Crutchfield, R. S., & Ballachey, E. L. (1962). *Individual in society: A textbook of social psychology*. McGraw-Hill.

Lavidas, K., Papadakis, S., Manesis, D., Grigoriadou, A. S., & Gialamas, V. (2022). The effects of social desirability on students' self-reports in two social contexts: Lectures vs. lectures and lab classes. *Information*, 13(10), 491.

Layne, T. (2015). Analysis of instructional impact on the running performance of university students. *Journal of Teaching, Research, and Media in Kinesiology*, 4, 1–14.

Li, F., Chen, J., & Baker, M. (2014). University students' attitudes toward physical education teaching. Journal of Teaching in Physical Education, 33(2), 186-212. https://doi.org/10.1123/jtpe.2013-0023

Li, G. C. (2014). *Analysis and research on factors affecting physical fitness development of middle school students in Chaoyang district, Beijing* (Master's thesis, Capital Institute of Physical Education).

Li, H., Cheong, J. P. G., & Hussain, B. (2023). The effect of a 12-week physical functional training-based physical education intervention on students' physical fitness—A quasi-experimental study. *International Journal of Environmental Research and Public Health*, 20(5), 3926. https://doi.org/10.3390/ijerph20053926

Li, P., Wang, W., Liu, H., Zhang, C., & Hastie, P. A. (2022). The impact of Sport Education on Chinese Physical Education majors' volleyball competence and knowledge. *Sustainability (Switzerland)*, 14(3), 1–11.

Li, X., Wang, Y., Li, X., et al. (2015). Study on the reliability and validity of the Chinese version of the Adolescent Physical Activity Questionnaire (PAQ-A). *Journal of Beijing University of Sport, 38*(5), 63–67.

Liang, W. (2024). A study on the implementation of the "Compulsory Education Physical Education and Health Curriculum Standards (2022 Edition)" in senior grades of primary schools in Guangzhou (Master's thesis, Guangzhou University). https://doi.org/10.27040/d.cnki.ggzdu.2024.000926

Liu, J., Liu, L., & Colditz, G. A. (2019). Optimal designs in three-level cluster randomized trials with a binary outcome. *Statistics in Medicine*, 38(20), 3733-3745. https://doi.org/10.1002/sim.8153

Liu, T., & Lipowski, M. (2021). Influence of cooperative learning intervention on the intrinsic motivation of physical education students—a meta-analysis within a limited range. *International Journal of Environmental Research and Public Health, 18*(6), 2989. https://doi.org/10.3390/ijerph18062989

Lubans, D. R., Leahy, A. A., Mavilidi, M. F., & Valkenborghs, S. R. (2021). Physical activity, fitness, and executive functions in youth: Effects, moderators, and mechanisms. In *Sensitive periods of brain development and preventive interventions* (pp. 103-130). Springer International Publishing.

Ludin, A. F. M. (2013). Effects of a 16-week, home-based, high-intensity, progressive resistance training program on glucose homeostasis, cardiovascular disease risk factors and functional capacity of older Malaysians with type 2 diabetes mellitus (Doctoral dissertation, Universiti Putra Malaysia).

Luna, P., Guerrero, J., Rodrigo-Ruiz, D., Losada, L., & Cejudo, J. (2020). Social competence and peer social acceptance: Evaluating effects of an educational intervention in adolescents. Frontiers in psychology, 11, 1305. https://doi.org/10.3389/fpsyg.2020.01305.

Luna, P., Guerrero, J., Rodrigo-Ruiz, D., Losada, L., & Cejudo, J. (2020). Social competence and peer social acceptance: Evaluating effects of an educational intervention in adolescents. Frontiers in psychology, 11, 1305. https://doi.org/10.3389/fpsyg.2020.01305

Ma, D. (2012). *Application of Sport Education Model in tennis teaching in colleges and universities* (Master's dissertation, Shandong University of Physical Education). CNKI.

Ma, H. (2020). Sport Education Model figure of eight big jump rope for elementary school teaching effects on the experimental study (Master's thesis, Eastern Shandong University).

MacPhail, A., Kirk, D., & Griffin, L. L. (2013). Throwing and catching as relational skills in game play: Situated learning in a modified game unit. *Journal of Teaching in Physical Education, 32*(3), 232-244.

Mallen, G., & Grenville-Cleave, B. (2017). Positive psychology and physical education in schools. In *Positive Psychology in Sport and Physical Activity* (pp. 182-192). Routledge.

Malshe, R., & Gunjal, S. B. (2021). Assessment of cardiovascular fitness amongst college students involved in sport activities and dance.

International Journal of Science and Healthcare Research, 6(2), 302-307.

Mann, T., Lamberts, R. P., & Lambert, M. I. (2013). Methods of prescribing relative exercise intensity: Physiological and practical considerations. *Sports Medicine, 43*, 613-625. https://doi.org/10.1007/s40279-013-0045-x

Marra, M., Sammarco, R., De Lorenzo, A., Iellamo, F., Siervo, M., Pietrobelli, A., ... & Contaldo, F. (2019). Assessment of body composition in health and disease using bioelectrical impedance analysis (BIA) and dual energy X-ray absorptiometry (DXA): A critical overview. *Contrast Media & Molecular Imaging*, 2019, 3548284.

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: Adapting and reinterpreting physical environments in the development of fraction concepts. *Cognitive Science, 29*(4), 587-625. https://doi.org/10.1207/s15516709cog0000_15

Matore, M. E. E. M., & Khairani, A. Z. (2015). Assessing the content validity of IKBAR using content validity ratio. *Australian Journal of Basic and Applied Sciences*, 9(7), 255-257.

McCaffrey, D. F., & Bell, R. M. (2006). Improved hypothesis testing for coefficients in generalized estimating equations with small samples of clusters. *Statistics in Medicine*, 25(23), 4081-4098. https://doi.org/10.1002/sim.2502

McGuire, W. J. (1989). The structure of individual attitudes and attitude systems. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 37-69). Erlbaum.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3), 276-282. https://hrcak.srce.hr/89395

Medrano-Ureña, M. D. R., Ortega-Ruiz, R., & Benítez-Sillero, J. D. (2020). Physical fitness, exercise self-efficacy, and quality of life in adulthood: a systematic review. *International Journal of Environmental Research and Public Health*, 17(17), 6343. https://doi.org/10.3390/ijerph17176343.

Melby, P. S., Elsborg, P., Nielsen, G., et al. (2021). 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, 21,* 1492. https://doi.org/10.1186/s12889-021-11343-1

Méndez-Giménez, A., del Pilar Mahedero-Navarrete, M., Puente-Maxera, F., & de Ojeda, D. M. (2022). Effects of the Sport Education model on adolescents' motivational, emotional, and well-being dimensions

during a school year. *European Physical Education Review, 28*(2), 380-396. https://doi.org/10.1177/1356336X211047866

Meng, C. (2022). Physical health measurement and analysis of the influencing factors of junior high school students in Guiyang city (Master's thesis, Guizhou Normal University). CNKI. https://doi.org/10.27114/d.cnki.gxndu.2022.001234

Menickelli, J., & Hastie, P. A. (2014). The impact of two curricular models on motivation, engagement, and achievement in physical education. *International Journal of Physical Education, Fitness and Sports, 3*(2), 33-42. https://doi.org/10.26524/1423

Metzler, M. (2017). *Instructional models for physical education* (3rd ed.). Scottsdale, AZ: Holcomb Hathaway Publishers.

Miller, D. K. (2006). *Measurement by the physical educator: Why and how* (5th ed.). McGraw-Hill Companies.

Miller, T., Herniman, J. J., Ricard, M. D., Cheatham, C. C., & Michael, T. J. (2012). The effects of a 6-week plyometric training program on agility. *Journal of Sports Science and Medicine, 5*(3), 459-465.

Mohsin, S. M. (1990). *Attitude: Concept, formation and change*. Wiley.

Morgans, R., Kweon, D., Ryan, B., Ju, W., Zmijewski, P., Oliveira, R., & Olthof, S. (2024). Playing position and match location affect the number of high-intensity efforts more than the quality of the opposition in elite football players. *Biology of Sport*, 41(3), 29-37.

Mosston, M., & Ashworth, S. (2008). *Teaching Physical Education* (1st online edition). Buckeystown, MD: Spectrum Institute for Teaching and Learning.

Murray, D. M., & Short, B. (1995). Intraclass correlation among measures related to alcohol use by young adults. *Journal of Studies on Alcohol, 56*(5), 681–694. https://doi.org/10.15288/jsa.1995.56.681

Murray, D. M., & Short, B. (1996). Intraclass correlation among measures related to alcohol use by school-aged adolescents: Estimates, correlates and applications in intervention studies. *Journal of Drug Education, 26*(3), 207–230. https://doi.org/10.2190/KBHP-6FRT-UOBN-VAUC

Murray, D. M., Rooney, B. L., Hannan, P. J., Peterson, A. V., Ary, D. V., Biglan, A., & Schinke, S. P. (1994). Intraclass correlation among common measures of adolescent smoking. *American Journal of Epidemiology, 140*(11), 1038–1050. https://doi.org/10.1093/oxfordjournals.aje.a117194

Narayan, K. A., & Khan, A. R. (2007). Body mass index and nutritional status of adults in two rural villages in Northern Malaysia. Malaysian Journal of Nutrition, 13(1), 9-17.

Neil-Sztramko, S. E., Caldwell, H., & Dobbins, M. (2021). School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. *Cochrane Database of Systematic Reviews,* (9). https://doi.org/10.1002/14651858.CD007651.pub3

Nikravan, M., & Zarei, A. (2019). Physical education method: Effects on physical fitness and competency of the students. *Annals of Applied Sport Science, 7*(4), 17-26.

Nikravan, M., & Zarei, A. (2019). Physical education method: Effects on physical fitness and competency of the students. *Annals of Applied Sport Science, 7*(4), 17–26.

Nishanbayevich, M. O. (2022). Football as a basic tool Physical Education of students. *Евразийский журнал академических исследований, 2*(2), 462-466.

Nokariya, S., Kotani, T., Sakuma, T., Iijima, Y., Okumura, T., Katogi, T., ... & Takaso, M. (2023). Trunk flexibility using a sit-and-reach test after surgery for adolescent idiopathic scoliosis. *Spine Deformity*, 11(2), 297-303.

Oppenheim, A. N. (1992). *Questionnaire design, interviewing and attitude measurement*. St. Martin's Press.

Opstoel, K., Chapelle, L., Prins, F. J., De Meester, A., Haerens, L., Van Tartwijk, J., & De Martelaer, K. (2020). Personal and social development in physical education and sports: A review study. *European Physical Education Review, 26*(4), 797-813. https://doi.org/10.1177/1356336X19882054

Orlić, A., Gromović, B., & Lakić, D. (2023). Influence of physical activity on emotional intelligence and self-regulation in adolescents. *Frontiers in Psychology, 14*, 1456191. https://doi.org/10.3389/fpsyg.2023.1456191

Orsso, C. E., Silva, M. I. B., Gonzalez, M. C., Rubin, D. A., Heymsfield, S. B., Prado, C. M., & Haqq, A. M. (2020). Assessment of body composition in pediatric overweight and obesity: A systematic review of the reliability and validity of common techniques. *Obesity Reviews*, 21(8), e13041.

OuYang, X. (2023). *Experimental research on the effect of HIIT on physical fitness of junior high school students* (Master's dissertation, Guangxi Normal University).

Özdemir, M. (2023). The effect of school-based physical activity on body composition and body mass index as indicators of obesity. *International Journal of Sports Engineering and Biotechnology*, 1(1), 25-28.

Pedretti, A., Mello, J. B., Gaya, A. R., Pedretti, A., & Gaya, A. C. A. (2020). Health-and skill-related physical fitness profile of Brazilian children and adolescents: A systematic review. *Revista Brasileira de Atividade Física & Saúde*, 25, 1-10.

Pérez-Ordás, R., Pozo, P., & Grao-Cruces, A. (2020). Effects on aggression and social responsibility by teaching personal and social responsibility during physical education. *Journal of Physical Education and Sport*, 20(4), 1832-1838.

Perlman, D. (2010). Change in affect and needs satisfaction for a motivated student within the sport education model. *Journal of Teaching in Physical Education, 29*(4), 433-445. https://doi.org/10.1123/jtpe.29.4.433

Petrušič, T., Trajković, N., & Bogataj, Š. (2022). Twelve-week game-based school intervention improves physical fitness in 12–14-year-old girls. *Frontiers in Public Health*, 10, 831424. https://doi.org/10.3389/fpubh.2022.831424

Petty, R. E., Wegener, D. T., & Fabrigar, L. R. (1997). Attitudes and attitude change. *Annual Review of Psychology, 48*, 609-647. https://doi.org/10.1146/annurev.psych.48.1.609

Pill, S., & Gambles, E. A. F. (2023). Teaching Games for Understanding and athlete-centred coaching. In S. Pill & E. A. F. Gambles (Eds.), *Teaching games and sport for understanding* (pp. 134-142). Routledge.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing and Health*, 29(5), 489-497. https://doi.org/10.1002/nur.20147

Premaratne, U. N., Sterne, J. A. C., Webb, J., & Burney, P. G. J. (1997). A general practice-based community intervention trial on the management of asthma. *Thorax, 52*(6), 1–6.

Pritchard, T., Hansen, A., Scarboro, S., & Melnic, I. (2015). Effectiveness of the Sport Education Fitness Model on fitness levels, knowledge, and physical activity. The Physical Educator, 72(4), 577–600. https://doi.org/10.18666/tpe-2015-v72-i4-6568

Pritchard, T., Hansen, A., Scarboro, S., & Melnic, I. (2015). Effectiveness of the sport education fitness model on fitness levels, knowledge, and

physical activity. *The Physical Educator, 72*(4), 577–600. https://doi.org/10.18666/tpe-2015-v72-i4-6568

Pritchard, T., Hawkins, A., Wiegand, R., & Metzler, J. N. (2008). Effects of two instructional approaches on skill development, knowledge, and game performance. *Measurement in Physical Education and Exercise Science*, 12(4), 219–236.

Quiñonero-Martínez, A. L., Cifo-Izquierdo, M. I., Sánchez-Alcaraz Martínez, B. J., & Gómez-Mármol, A. (2023). Effect of the hybridization of social and personal responsibility model and sport education model on physical fitness status and physical activity practice. *Frontiers in Psychology, 14*, 1273513. https://doi.org/10.3389/fpsyg.2023.1273513

Reddy, S., & LaBarbera, P. (1985). Hierarchical models of attitude. *Multivariate Behavioral Research, 20*(4), 451-471. https://doi.org/10.1207/s15327906mbr2004_4

Ren, Z. (2024). A study on the implementation of the "Compulsory Education Physical Education and Health Curriculum Standards (2022)" in Chishui City, Guizhou (Master's thesis, Guizhou Normal University). https://doi.org/10.27048/d.cnki.ggzsu.2024.000139

Rivera-Pérez, S., Fernandez-Rio, J., & Iglesias Gallego, D. (2021). Effects of an 8-week cooperative learning intervention on physical education students' task and self-approach goals, and emotional intelligence. *International Journal of Environmental Research and Public Health, 18*(1), 61. https://doi.org/10.3390/ijerph18010061

Robinson, L. E., Stodden, D. F., Barnett, L. M., Lopes, V. P., Logan, S. W., Rodrigues, L. P., & D'Hondt, E. (2014). Motor competence and its effect on positive developmental trajectories of health. *Sports Medicine, 45*(9), 1273-1284. https://doi.org/10.1007/s40279-014-0223-6

Rosado, A., & Mesquita, I. (2009). *Melhorar a aprendizagem optimizando a instrução*. In A. Rosado & I. Mesquita (Eds.), *Pedagogia do Desporto* (pp. 69-130). Lisboa: Edições FMH - UTL.

Rosenberg, M. J. (1960). A structural theory of attitude dynamics. *Public Opinion Quarterly, 24*(2), 319-340. https://doi.org/10.1086/266951

Saavedra, J. M., Torres, S., Caro, B., Escalante, Y., Cruz, E. L. A., Duran, M. J., & Rodríguez, F. A. (2008). Relationship between health-related fitness and educational and income levels in Spanish women. Public Health, 122(8), 794-800. https://doi.org/10.1016/j.puhe.2008.03.004

Schwarz, N., & Bohner, G. (2001). The construction of attitudes. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Intraindividual processes* (pp. 436–457). Blackwell.

Seeman, M. (1993). A historical perspective on attitude research. In D. Krebs & P. Schmidt (Eds.), *New directions in attitude measurement* (pp. 3-20). de Gruyter.

Segovia, Y., & Gutiérrez, D. (2020). Effect of a game-based high intensity interval training program on body composition in primary education: Comparison of the Sport Education model and traditional methodology. Journal of Physical Education and Sport, 20(2), 791-799. https://doi.org/10.7752/jpes.2020.02115

Shalar, O., Strykalenko, Y., Huzar, V., Voloshinov, S., Homenko, V., & Svirida, V. (2021). Physical fitness assessment of young football players using an integrated approach. *Journal of Physical Education and Sport*, 21(5), 2674-2680.

Shao, Z., Shen, Q., Yao, B., Mao, C., Chen, L. N., Zhang, H., ... & Shen, H. (2022). Identification and mechanism of G protein-biased ligands for chemokine receptor CCR1. Nature Chemical Biology, 18(3), 264-271. https://doi.org/10.1038/s41589-021-00888-7

Shapiro, L. (2019). *Embodied cognition* (2nd ed.). Routledge.

Shen, H., Yan, J., Hong, J. T., Clark, C., Yang, X. N., Liu, Y., & Chen, S. T. (2020). Prevalence of physical activity and sedentary behavior among Chinese children and adolescents: Variations, gaps, and recommendations. *International Journal of Environmental Research and Public Health, 17*(9), 3066. https://doi.org/10.3390/ijerph17093066

Siedentop, D., & Van der Mars, H. (2022). *Introduction to physical education, fitness, and sport*. Human Kinetics.

Siedentop, D., Hastie, P., & van der Mars, H. (2011). *Complete guide to sport education*. Champaign, IL: Human Kinetics.

Siedentop, D., Hastie, P., & van der Mars, H. (2020). *Complete guide to sport education* (3rd ed.). Champaign, IL: Human Kinetics.

Silverman, S. (2017). Attitude research in physical education: A review. *Journal of Teaching in Physical Education*, 36(3), 303-312.

Silverman, S. (2017). Attitude research in physical education: A review. *Journal of Teaching in Physical Education, 36*(3), 303-312. https://doi.org/10.1123/jtpe.2017-0037

Silverman, S., & Subramaniam, P. R. (1999). Student attitude toward physical education and physical activity: A review of measurement issues and outcomes. Journal of Teaching in Physical Education, 19(1), 97-125. https://doi.org/10.1123/jtpe.19.1.97

Smedegaard, S., Christiansen, L. B., Lund-Cramer, P., Bredahl, T., & Skovgaard, T. (2016). Improving the well-being of children and youths:

A randomized multicomponent, school-based, physical activity intervention. *BMC Public Health, 16*(1), 1–11. https://doi.org/10.1186/s12889-016-2768-8

Spittle, M., & Byrne, K. (2009). The influence of Sport Education on student motivation in physical education. *Physical Education & Sport Pedagogy, 14*(3), 253-266.

Stankovic, M., Djordjevic, D., Katanic, B., Petrovic, A., Jelaska, I., Bjelica, D., & Mekic, A. (2023). Speed and agility training in female soccer players: A systematic review. *Sport Mont*, 21(3), 117-122.

Stantcheva, S. (2023). How to run surveys: A guide to creating your own identifying variation and revealing the invisible. *Annual Review of Economics*, 15(1), 205-234. https://doi.org/10.1146/annurev-economics-091622-010157.

Stastny, P., Musalek, M., Roczniok, R., Cleather, D., Novak, D., & Vagner, M. (2023). Testing distance characteristics and reference values for ice-hockey straight sprint speed and acceleration: A systematic review and meta-analyses. *Biology of Sport*, 40(3), 899-918.

Stratton, G. (1996). Children's heart rates during physical education lessons: A review. *Pediatric Exercise Science*, 8(3), 215-233.

Su Yuanyuan, Chang Hongjun, Qian Run, & Zhang Hongyue. (2023). A study on children's outdoor physical activity behavior preference in urban relocation communities: A case study of Luyiju Community in Hefei City. *Journal of Southwest University (Natural Science Edition)*, 12, 178-192. https://doi.org/10.13718/j.carolcarrollnki.xdzk.2023.12.018

Subramaniam, P. R., & Silverman, S. (2000). Validation of scores from an instrument assessing student attitude toward physical education. *Measurement in Physical Education and Exercise Science*, 4(1), 29-43.

Subramaniam, P. R., & Silverman, S. (2007). Middle school students' attitudes toward physical education. *Teaching and Teacher Education*, 23(5), 602-611.

Sun, K. (2018). *Research on the application of sport education curriculum in middle school physical education* (Master's thesis). Shaanxi Normal University.

Sun, Q. (2013). *Research on the application of sport education curriculum in middle school physical education* (Master's thesis). Jiangxi Normal University.

Sun, S., & Duan, B. (2020). Application and effect of BMI in evaluating the physical health of adolescents. China Journal of School Health, 41(2), 247-251.

Suyati, E. S., Sonedi, S., Bulkani, B., Fatchurahman, M., Nurbudiyani, I., & Setiawan, M. A. (2022). The relationship of physical fitness and social-economic status and students' learning achievement. Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación, (46), 494-500. https://doi.org/10.47197/retos.v46i46.95151

Tendinha, R., Alves, M. D., Freitas, T., Appleton, G., Gonçalves, L., Ihle, A., ... & Marques, A. (2021). Impact of sports education model in physical education on students' motivation: A systematic review. Children, 8(7), 588. https://doi.org/10.3390/children8070588

Tinôco, M. A., Nascimento, M. D. M., Marques, A., Gouveia, É. R., Miguel, S., Tinôco, F., & Ihle, A. (2023). The relationship between physical fitness and cognitive functions in older people: A systematic review. Sustainability, 15(23), 16314. https://doi.org/10.3390/su152316314

Titans Football Academy. (2023). Benefits of football: Improve health and fitness. *The Titans Football Academy.* https://thetitansfa.com

Trevino, R. P., Fogt, D. L., Wyatt, T. J., Leal-Vasquez, L., Sosa, E., & Woods, C. (2008). Diabetes risk, low fitness, and energy insufficiency levels among children from poor families. Journal of the American Dietetic Association, 108(11), 1846-1853. https://doi.org/10.1016/j.jada.2008.08.013

Trist, L., Wallhead, A., Garn, C. A., & Vidoni, C. (2014). Effect of a sport education program on motivation for physical education and leisure-time physical activity. *Research Quarterly for Exercise and Sport*, 85(4), 478-487.

Vadlin, S., Åslund, C., & Nilsson, K. W. (2015). Development and content validity of a screening instrument for gaming addiction in adolescents: The Gaming Addiction Identification Test (GAIT). *Scandinavian Journal of Psychology*, 56(4), 458-466. https://doi.org/10.1111/sjop.12196

Van Den Tillaar, R., & Ball, N. (2020). Push-ups are able to predict the bench press 1-RM and constitute an alternative for measuring maximum upper body strength based on load-velocity relationships. *Journal of Human Kinetics*, 73(1), 7-18.

Varela, F. J., Thompson, E., & Rosch, E. (2017). *The embodied mind, revised edition: Cognitive science and human experience*. MIT Press.

Wahl-Alexander, Z., & Chomentowski, P. (2018). Impact of a university physical conditioning sport education season on students' fitness levels. *Health Education Journal,* 77(7), 828–836. https://doi.org/10.1177/0017896918776340

Wahl-Alexander, Z., & Chomentowski, P. (2018). Impact of a university physical conditioning sport education season on students' fitness

levels. Health Education Journal, 77(7), 828–836. https://doi.org/10.1177/0017896918776340

Wahl-Alexander, Z., & Chomentowski, P. (2018). Impact of a university physical conditioning sport education season on students' fitness levels. *Health Education Journal,* 77(7), 828–836. https://doi.org/10.1177/0017896918776340

Wallhead, T. L., & Ntoumanis, N. (2004). Effects of a sport education intervention on students' motivational responses in physical education. *Journal of Teaching in Physical Education, 23*(1), 4-18.

Wallhead, T. L., Garn, A. C., & Vidoni, C. (2014). Effect of a sport education program on motivation for physical education and leisure-time physical activity. *Research Quarterly for Exercise and Sport, 85*(4), 478-487.

Wang, H., Chen, C., & Jiang, Y. (2023). Construction of a volleyball skill learning evaluation model for junior high school students: Based on the "Compulsory Education Physical Education and Health Curriculum Standards (2022 Edition)." *Sports Research and Education*, 38(5), 78-88. https://doi.org/10.16207/j.cnki.2095-235x.2023.05.006

Wang, Q., & Zhang, T. (2021). Physical exercise and its impact on students' physical health and cognitive development. *BMC Public Health*, 21(1), 1120-1130.

Wang, X. (2024). The impact of sports game teaching method on fifth-grade students' interest in physical education and physical fitness (Master's thesis, Nanjing Sport Institute). https://doi.org/10.27247/d.cnki.gnjtc.2024.000118

Ward, J. K., Hastie, P. A., Wadsworth, D. D., Foote, S., Brock, S. J., & Hollett, N. (2017). A sport education fitness season's impact on students' fitness levels, knowledge, and in-class physical activity. Research Quarterly for Exercise and Sport, 88(3), 346-351. https://doi.org/10.1080/02701367.2017.1321100

Ward, P., & Griggs, G. (2011). Principles of instructional design for sport education. *Journal of Teaching in Physical Education, 30*(3), 269-282. https://doi.org/10.1123/jtpe.30.3.269

Wibisana, M. I. N., & Royana, I. F. (2023). A randomized controlled trial on the effectiveness of FITT principle in increasing VO2 max. *Musamus Journal of Physical Education and Sport (MJPES), 6*(1), 295-304. https://doi.org/10.35139/mjpes.v6i1.122

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology, 4*, 58.

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology, 4*, 58. https://doi.org/10.3389/fpsyg.2013.00058

Wisløff, U., Castagna, C., Helgerud, J., Jones, R., & Hoff, J. (2004). Strong correlation of maximal squat strength with sprint performance and vertical jump height in elite soccer players. *British Journal of Sports Medicine, 38*(3), 285–288. https://doi.org/10.1136/bjsm.2002.002071

World Health Organization. (2021, February 17). WHO reviews effect of physical activity on enhancing academic achievement at school. *WHO News.* https://www.who.int

Wu, W. (2008). Experimental research on Sport Education Model in physical education teaching in colleges and universities. Journal of Beijing Sport University, 31(12), 35-38.

Xiao, Q. (2012). Analysis on the influencing factors of physical health of left-behind middle school students in rural minority areas: A case study of Tongren City (Master's thesis, Huazhong University of Science and Technology). CNKI. https://doi.org/10.27114/d.cnki.ghzju.2012.000051

Xiao, W. (2022). Differences in selected fitness, tennis skills, and functional movement following functional movement training program among young male Chinese tennis players (Doctoral thesis, Universiti Putra Malaysia).

Xie Xiangmei. (2006). *Study on the attitude of PE learning in middle school students*. Nanjing Normal University.

Xie, H. (2012). Influence of physical health belief on intention and habit of physical exercise behavior of college students. [Master's thesis, Beijing Sport University]. China National Knowledge Infrastructure (CNKI).

Xu, C., Rong, G., & Shuaijing, X. (2019). Impact of a sport education season on students' table tennis skills and attitudes in China's high school. *International Journal of Information and Education Technology, 9*(11). https://doi.org/10.18178/ijiet.2019.9.11.1311

Xu, X., He, F. Z., & Wang, J. P. (2021). A comparative study of the new era physical education curriculum and the Sport Education Model. *University*, (31), 25-27.

Xu, X., He, F., & Wang, J. (2021). A comparative study of physical education curriculum and Sport Education Model in the new era. *University*, 31, 25-27.

Xue, E., & Li, J. (2023). Policy analysis of educational resource allocation in Hebei province. In *Coordinated Education Development Policy in*

China. Exploring Education Policy in a Globalized World: Concepts, Contexts, and Practices. Springer, Singapore. https://doi.org/10.1007/978-981-19-7931-6_3

Yang, L. (2022). *Research on the interaction of sense of meaning of life, body cognition, and physical exercise behavior in junior middle school students* (Master's dissertation, Tianjin Institute of Sport).

Yin, J. (2019). Experimental study of sport education mode in the teaching of primary school football (Master's thesis, Hunan University of Science and Technology). CNKI. https://doi.org/10.27738/d.cnki.ghnkd.2019000133

Young, L., O'Connor, J., Alfrey, L., & Penney, D. (2021). Assessing physical literacy in health and physical education. *Curriculum Studies in Health and Physical Education, 12*(2), 156-179. https://doi.org/10.1080/25742981.2020.1810582

Zajonc, R. B., & Markus, H. (1982). Affective and cognitive factors in preferences. *Journal of Consumer Research, 9*, 123-131. https://doi.org/10.1086/208905

Zalleg, D., Dhahbi, A. B., Dhahbi, W., Sellami, M., Padulo, J., Souaifi, M., ... & Chamari, K. (2020). Explosive push-ups: From popular simple exercises to valid tests for upper-body power. *The Journal of Strength & Conditioning Research*, 34(10), 2877-2885.

Zhang, J., Soh, K. G., Anuar, M. A. B. M., Lin, X., & Nasiruddin, N. J. M. (2023). The Effect of the Sport Education Model on sports ability: A Critical Systematic Review. *Revista de Psicología del Deporte (Journal of Sport Psychology), 32*(3), 339-354.

Zhang, J., Xiao, W., Soh, K. G., Yao, G., Anuar, M. A. B. M., Bai, X., & Bao, L. (2024). The effect of the Sport Education Model in physical education on student learning attitude: a systematic review. BMC public health, 24(1), 949. https://doi.org/10.1186/s12889-024-18243-0

Zhang, N., & Su, R. (2020). Longitudinal research on the application of the Sport Education Model to college physical education. *Revista de Cercetare si Interventie Sociala*, 70. https://doi.org/10.33788/rcis.70.21.

Zhang, S. (2020). Positioning according to the origin, interest, and truth—Decoding the phenomenon that middle school students prefer sports rather than physical education. *Sports Teacher and Friends, 43*(05), 68-70.

Zhang, Y. (2022). The impact of fancy rope skipping exercises on the physical fitness and mental health of seventh-grade students (Master's thesis, Hebei Normal University). https://doi.org/10.27110/d.cnki.ghsfu.2022.000287

Zhen-Rong, J. I. A. (2021). Effects of teaching games for understanding integrated sport education model on college students' football cognitive performance and motor skills. *Revista de Cercetare şi Intervenţie Socială, (72),* 274-287. https://doi.org/10.33788/rcis.72.19

Zhu, Y. (2022). The impact of achievement motivation on middle school students' learning engagement: The role of time management tendency and self-control and educational implications (Master's thesis, Henan University). CNKI. https://doi.org/10.27114/d.cnki.ghnau.2022.001803

Zhuo Zhichao. (2021). A comparative study of three sports modes. *Stationery & Science & Technology*, (04), 123-125.

APPENDICES

Appendix A

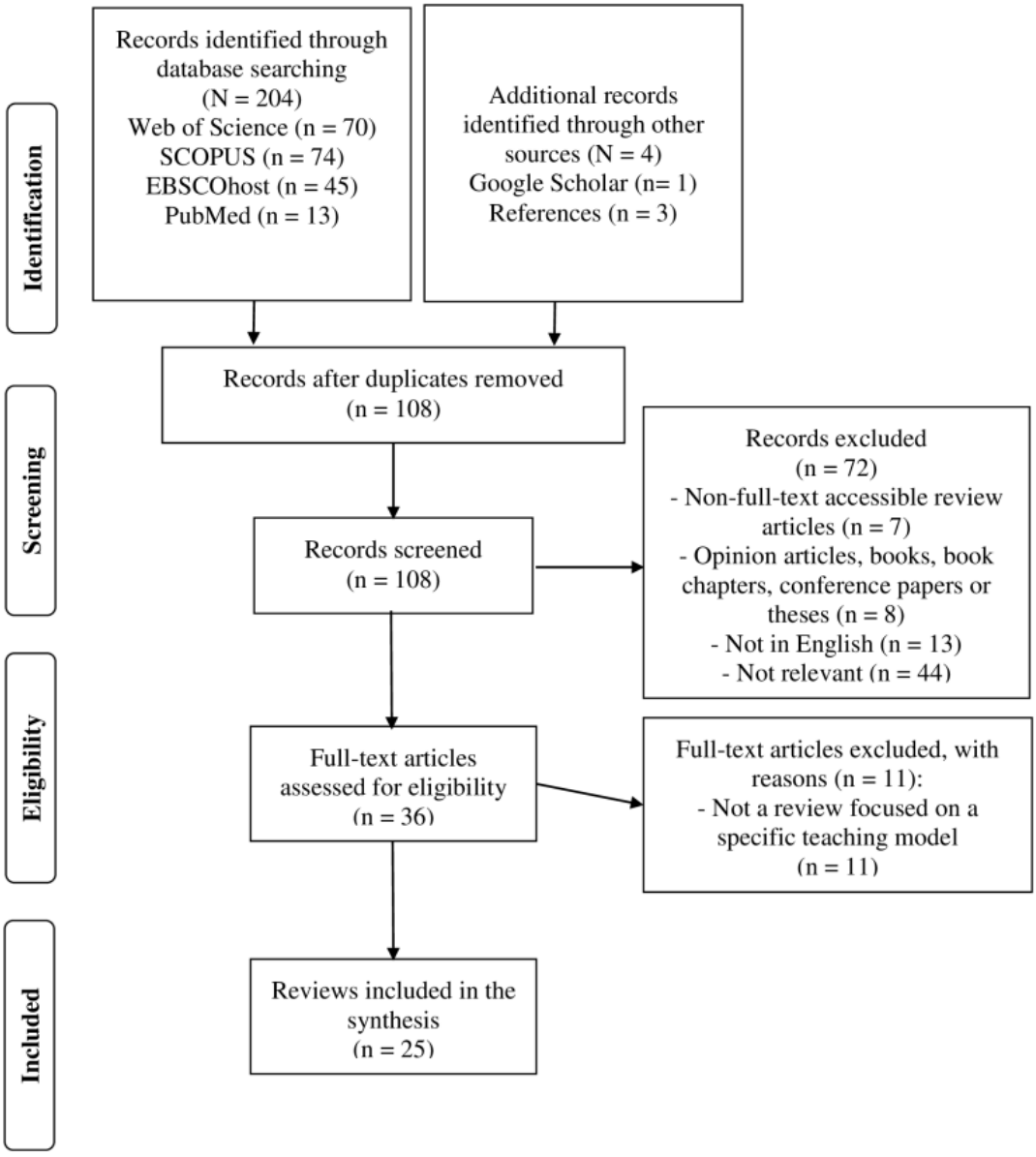


Figure 1 : Flow-Chart of Reviews' Selection Procedure

Table 1 : Detailed Search Strategy

Search database	Keyword	Number
Web of Science	((TS=("pedagogical model*" OR "curriculum model*" OR "model*-based approach*" OR "instructional model*," OR "game-centered approach*" OR "sport education" OR "teaching games for understanding" OR "teaching for personal and social responsibility" OR "cooperative learning" OR "tactical games approach" OR "game sense" OR "tactical decision learning model" OR "developmental game stages model")) AND TS=("physical education")) AND TI=(review)	n = 70
SCOPUS	TITLE-ABS-KEY ("pedagogical model*" OR "curriculum model*" OR "model*-based approach*" OR "instructional model*," OR "game-centred approach*" OR "sport education" OR "teaching games for understanding" OR "teaching for personal and social responsibility" OR "cooperative learning" OR "tactical games approach" OR "game sense" OR "tactical decision learning model" OR "developmental game stages model") AND TITLE-ABS-KEY-AUTH ("physical education") AND TITLE-ABS-KEY-AUTH ("review")	n = 74
EBSCOhost (CINAHL Plus with Full Text, SPORTDiscus with Full Text,PsycINFO,and PsycARTICLES)	AB ("pedagogical model*" OR "curriculum model*" OR "model*-based approach*" OR "instructional model*," OR "game-centered approach*" OR "sport education" OR "teaching games for understanding" OR "teaching for personal and social responsibility" OR "cooperative learning" OR "tactical games approach" OR "game sense" OR "tactical decision learning model" OR "developmental game stages model") AND AB "physical education" AND TI review	n = 45
PubMed	(("pedagogical model*"[Title/Abstract] OR "curriculum model*"[Title/Abstract] OR "model*-based approach*"[Title/Abstract] OR "instructional model*,"[Title/Abstract] OR "game-centred approach*"[Title/Abstract] OR "sport education"[Title/Abstract] OR "teaching games for understanding"[Title/Abstract] OR "teaching for personal and social responsibility"[Title/Abstract] OR "cooperative learning"[Title/Abstract] OR "tactical games approach"[Title/Abstract] OR "game sense"[Title/Abstract] OR "tactical decision learning model"[Title/Abstract] OR "developmental game stages model"[Title/Abstract]) AND ("physical education"[Title/Abstract])) AND (review[Title/Abstract]) AND (review[Title/Abstract])	n = 13

Table S2 : Summary of the Results Reported

Source of Reviews	Pedagogical model	Purpose of Reviews	Mentioned theory	Educational outcomes	Implementation requires attention	Research gaps
Wallhead and O'Sullivan 2005 [31]	SEM (28 studies)	To summarize the impact of the SEM on student learning in physical education.	Play Theory	SEM improves student engagement and social skills like responsibility and cooperation.	Issues with student leadership may hinder content development and fairness.	Further study is needed on peer interactions and validating SE's cultural benefits.
Hastie et al. 2011 [10]	SEM (38 studies)	To update research on the SEM since 2005, identify new trends, and assess progress on previous research gaps.	Theory of motivation	SEM research has expanded to diverse contexts and addresses various physical education standards.	Issues persist in peer instruction and applying school-based learning to community sports.	Further studies are needed on peer instruction and the transition from school to community sports.
Harvey et al. 2014 [32]	SEM (21studies)	To explore how the SEM fosters sports literacy and ethical behaviours in youth.	Peace Theory	SEM includes concepts like inclusion, responsibility, and social justice, but teaching ethical conduct is essential	Emphasizing active teaching methods is needed for ethical behavior development.	Further focus is needed on practical applications like ethical contracts, sports panels, modified games, and reward systems.
Araújo et al. 2014 [11]	SEM (23 studies)	To assess what is known about student learning during participation in SEM.	Not mentioned	SEM has shown benefits in students' personal and social development, but learning outcomes vary.	Issues include unequal learning opportunities and the need for better control over the teaching-learning process.	Future studies should focus on student-coach preparation, hybrid approaches, and the role of gender and skill level in learning outcomes.
Chu and Zhang 2018 [12]	SEM (18 studies)	To review literature on motivational	Self-determination Theory and	Sport Education positively	Long-term follow-up studies and	More research is required to explore the

		processes in Sport Education using self-determination and achievement goal theories.	Achievement Goal Theory	influences motivation across genders, grade levels, and sports.	inclusion of teacher participants are needed.	motivational impact in varied school settings over time.
Evangelio et al. 2018 [13]	SEM (38 studies)	This study reviews recent implementations of the SEM for grades 1 to 12 and highlights emerging research trends.	Not mentioned	SEM has expanded to include all learning domains: physical, social, cognitive, and affective.	Specific teacher training is essential for effective SEM implementation.	Future research should focus on motor skill development, tactical knowledge, and values within SEM.
Bessa et al. 2019 [14]	SEM (51 studies)	This systematic review examines the impact of SEM versus Traditional Teaching (TT) on students' learning outcomes in Physical Education.	Not mentioned	Results indicate that SE leads to superior outcomes in developing personal, social, motor, and cognitive skills compared to TT.	Future research should focus on using diverse research methods, longer interventions, and ensuring model fidelity to enhance educational effectiveness.	The current studies predominantly involve quasi-experimental designs and high school students, highlighting the need for varied methodologies and broader participant demographics in future investigations.
Bessa et al. 2021[15]	SEM (28 studies)	This review compares the effects of the SEM and Traditional Teaching (TT) on learning outcomes in Physical Education.	Not mentioned	Research concerning the impact of SE on students' personal and social development has shown unequivocal results.	Few studies established the fidelity of the model implementation.	Future studies should use varied methodologies and longer teaching units for thorough analysis.
Tendinha et al. 2021 [16]	SEM (14 studies)	This study reviews the impact of the SEM on students'	Self-determination theory	The SEM significantly enhances	Incorporating the SEM in PE programs can	Further research is needed to understand SEM's long-term

		motivation in Physical Education (PE).		motivation, especially autonomy and enjoyment in PE.	effectively promote intrinsic motivation for physical activity.	effects on motivation in various educational contexts.
Manninen and Campbell 2022 [17]	SEM (25 studies)	This study synthesizes the effects of the SEM on need satisfaction, intrinsic motivation, and prosocial attitudes in physical education (PE).	Self-determination theory	The SEM positively impacts autonomy, competence, relatedness, intrinsic motivation, and prosocial attitudes.	Future SEM implementations should ensure high fidelity and transparency in reporting.	More high-quality trials are needed to explore broader outcomes, including novelty, negative effects, and essential behaviors like physical activity.
Zhang et al. 2023 [33]	SEM (12 studies)	To compare the effects of the SEM and Traditional Teaching (TT) on students' athletic skills.	Not mentioned	SEM generally outperforms TT in improving students' athletic abilities.	Key factors include following unit guidelines, proper intervention planning, and ensuring model fidelity.	More research is needed on SEM's impact in primary and high school settings to improve generalizability.
Ma et al. 2023 [34]	SEM (10 studies)	To evaluate the effectiveness of the SEM on content knowledge, technical skills, and game performance.	Not mentioned	SEM significantly improves content knowledge technical skills, and game performance.	Variations in learning stages and game performance measurement methods, rather than frequency and duration, impact SEM outcomes.	More high-quality studies across diverse sports are needed to better understand SEM's full potential.
Zhang et al. 2024 [2]	SEM (13 studies)	To analyze the impact of the SEM on students' attitudes toward physical education learning.	Attitude Theory	The SEM shows significant advantages over Traditional Teaching (TT) in improving students' cognitive and	Effective implementation requires better control of inter-group atmosphere, more frequent	There is a need for more data on junior high school students, gender differences, and improved evaluation methods for learning attitudes.

				affective aspects.	interventions, model fidelity checks, and appropriate measurement tools.	
Casey and Goodyear 2015 [21]	CL (27 studies)	To examine research on Cooperative Learning's impact on physical, cognitive, social, and affective learning in physical education.	Social Interdependence Theory	Learning is evident across physical, cognitive, and social domains, with affective learning less documented.	More focus is needed on long-term learning outcomes beyond short-term instructional units.	Further research is needed on affective learning and sustained impacts over longer periods.
Fernández-Espínola et al. 2020 [22]	CL (5 studies)	To evaluate the impact of cooperative learning interventions on intrinsic motivation in physical education and perform a meta-analysis of their effect size.	self-determination Theory; Achievement Goal Theory	Collaborative learning interventions are effective teaching strategies for enhancing intrinsic motivation among physical education students.	Students' age may influence their execution of the curriculum, as the strategies implemented in the curriculum can be quite complex.	The findings suggest caution due to high heterogeneity and low-quality evidence, emphasizing the need for more rigorous research.
Bores-García et al. 2021 [35]	CL (15 studies)	The review analyzes Cooperative Learning research in Physical Education from 2014 to 2019 across educational stages.	Not mentioned	Social learning outcomes like motivation and teacher-student interaction are emphasized, especially in secondary education.	The review highlights the need for rigorous implementation, criticizing the short and fragmented nature of many programs.	There is limited focus on body expression and a need for longer-term studies in Cooperative Learning.
Zach et al. 2023 [23]	CL (44 studies)	This article examines the applicability of field studies on cooperative learning in physical education	Not mentioned	Cooperative teaching strategies are partially linked to positive outcomes in social,	The intervention strategies are not clearly defined, lacking replicability.	Additional studies are required to fully understand the impact of cooperative learning in physical

		for optimizing teaching strategies.		physical, affective, and cognitive domains.		education.
Iglesias et al. 2023 [24]	CL (48 studies)	This note aims to provide an overview of research on cooperative learning in physical education.	Not mentioned	The findings highlight a significant focus on intervention studies, indicating a growing interest in cooperative learning strategies.	While research is expanding, it shows disparities in geographical distribution and the representation of women researchers.	The review indicates a need for more balanced contributions from different regions and increased participation of female authors in future studies.
Suryadi et al. 2024 [36]	CL (11 studies)	This review updates the literature on cooperative learning in physical education from 2019 to 2024.	Not mentioned	Cooperative learning enhances students' motivation, participation, social skills, and learning achievements.	Adapting cooperative learning strategies to specific learning goals and student needs is crucial for maximizing benefits.	Further studies, including global reviews and bibliometric analyses, are needed to deepen understanding of cooperative learning's impact, considering contemporary social values.
Zhou & Colomer 2024 [37]	CL (50 studies)	This review explores the impact of cooperative learning (CL) on student development and cultural diversity.	Social Constructivism Theory	CL enhances motor skills, social skills, and cognitive abilities, while reducing negative emotions.	CL's effectiveness varies by age and cultural context, with less focus on cognitive outcomes.	More studies are needed on CL's cognitive impact and its application across diverse settings.
Barba-Martín et al. 2020 [19]	TGFU (12 studies)	This systematic review analyzes Cooperative Learning research in Physical Education from 2014 to 2019.	Not mentioned	The TGFU model is suitable for promoting students' cognitive, motor, social, and emotional	Rigorous implementation of Cooperative Learning interventions in schools is essential.	Future studies should address longer interventions and better representation of body expression.

				development in teaching.		
Morales-Belando et al. 2022 [38]	TGFU (20 studies)	This review explores how TGFU researchers report their interventions and their connection to teaching-learning implementation features and learner outcomes.	The Theory of Practice Architecture	The study found inconsistent reporting of implementation features, complicating conclusions about their impact on student learning outcomes.	Future TGFU interventions should ensure context-based lesson design, clearly define intervention parameters, prioritize teacher training, validate lessons before implementation, and standardize outcome measurements.	There is a lack of comprehensive studies incorporating all teaching-learning implementation features, indicating a need for systematic research in TGFU interventions.
Ortiz et al. 2023 [20]	TGFU (13 studies)	This review examines the effects of the Teaching Games for Understanding (TGFU) intervention on game performance and psychosocial variables.	Not mentioned	TGFU significantly improves decision-making skills and moderately impacts motivational climate and task orientation.	The effects of TGFU on perceived competence, enjoyment, and game knowledge are unclear.	Further studies are needed to explore TGFU's influence on psychosocial variables and game knowledge.
Pozo et al. 2018 [25]	TPSR (22 studies)	This study reviews research on the Teaching Personal and Social Responsibility (TPSR) model in physical education.	Self-Determination Theory	The educational outcomes are categorized by their impact on teachers, implementation features, and student	Attention is needed on program implementation features within physical education.	There are gaps in research, particularly for longer studies, follow-up data, quantitative methods, and larger sample sizes.

				participants.		
Carrellán et al. 2023 [39]	TPSR (22 studies)	This review analyzes the impact of TPSR in Physical Education on student behavior.	Not mentioned	TPSR interventions improve Basic Psychological Needs, behavior, emotional understanding, and educational values.	Effective application of TPSR in PE can foster responsibility, offering valuable guidance for teachers and coaches.	More studies are needed to further explore TPSR long-term effects and optimize its implementation in various contexts.

Note: SEM: Sport Education Model; CL:Cooperative Learning; TGFU: Teaching Games for Understanding; TPSR: Teaching Personal and Social Responsibility

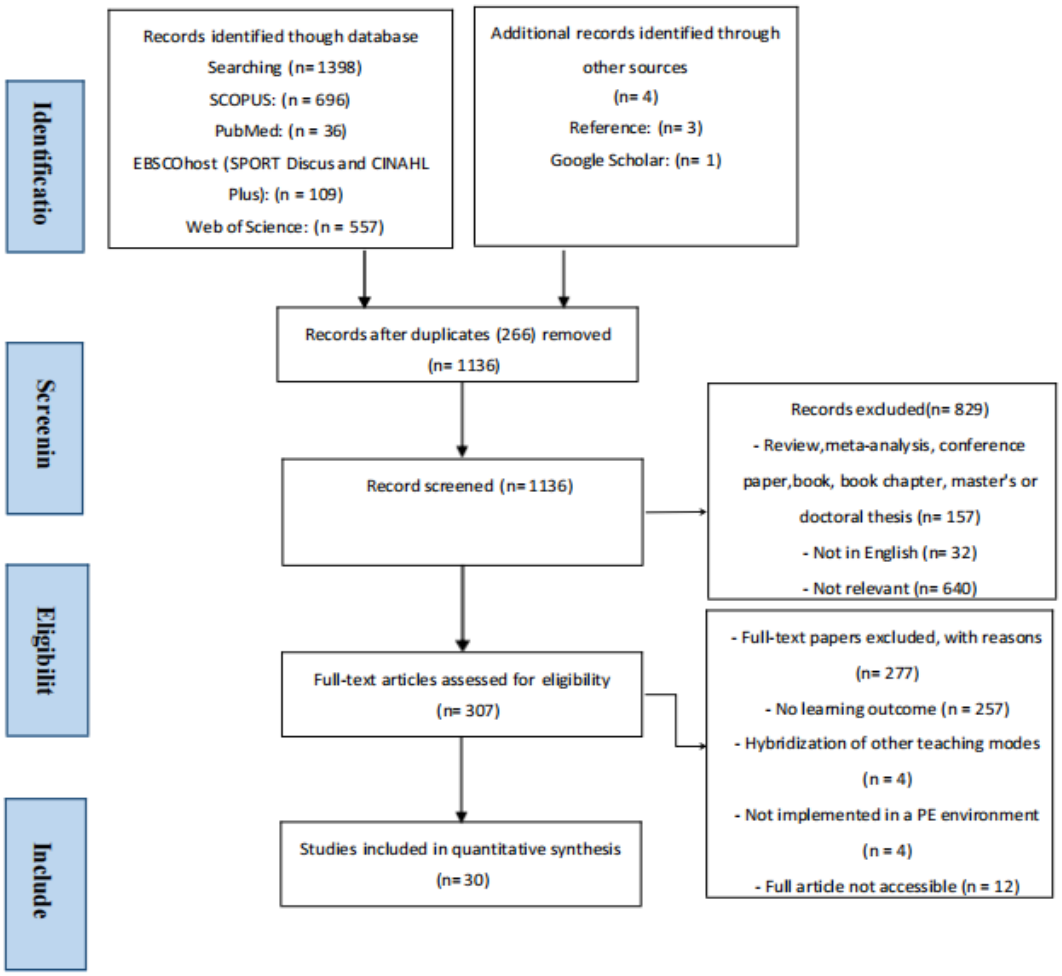


Figure 2 : PRISMA Flow Diagram

Table 4 : Detailed Search Strategy

Search database	keyword
	TITLE-ABS-KEY ("Sport Education") AND ("pedagogical model" OR "curriculum model" OR "instructional model" OR "physical education" OR "hybrid*" OR "Integrate*" OR "combine*")
Web of Science	n = 557
Scopus	n = 696
PubMed	n = 36
EBSCO host	n = 109

Table 5 : Data Extraction Included in the Study

Sources	Purpose	Sample Characteristics	Hybrid	Length of Unit /Content	Data Collection	Study Design & Analysis	Learning Outcomes
Araújo et al. (2016) Portugal	Analyzing students' improvements in game performance during a hybrid SEM-SGA approach	17 grade 7 students, consisting of 7 girls and 10 boys, with an average age of 11.8 years	SEM - SGA	25 lessons volleyball	Video Observation, Game Performance Assessment Tool (GPAI)	Quantitative: analysis of variance (ANOVA)	Motor, cognitive
Araújo et al. (2017) Portugal	Examining the impact on student coaches' pedagogical content knowledge (PCK)	21 students (11 males and 10 females); mean age 12.0	SEM - SGA	20-25 lessons volleyball	Video observation, field notes, interviews	Qualitative: thematic analysis method	Cognitive
Araújo et al. (2019) Portugal	Analysis of student competition performance in three SEM-SGA seasons	18 grade 7 students (8 female and 10 male) 11-13 years old	SEM - SGA	20-25 lessons volleyball	Video observation, the Game Performance Assessment Instrument (GPAI)	Quantitative: hierarchical linear model	Motor
Araújo et al. (2020) Portugal	Investigating students' tactical understanding	96 students (40 boys and 53 girls, aged from 16-18 years-old)	SEM - SGA	24 lessons volleyball	Video-based test	Quantitative: Mann-Whitney test	Motor, cognitive
Arikan (2020) Turkey	Analysis of SEM-SEL on students' emotional intelligence levels	166 students, aged from 15-17 years-old)	SEM - SEL	16 lessons volleyball	The Schutte Emotional Intelligence Scale (SEI-S)	Quantitative: Covariance Analysis	Social
Buendía et al. (2021) Spain	To compare the sportsmanship and enjoyment between the SEM-TGFU and PSRM	85 teenagers, 39 girls, and 46 boys ages 16.42 ± 0.50	SEM - TGFU	10 lessons ultimate frisbee	The SSI questionnaires adapted to Physical Education	Quantitative: Wilcoxon's test	Affective

					(SSI-PE); the Multidimensional Sportsmanship Orientations Scale (MSOS)		
Evangelio et al. (2021) Spain	Explore students' perceptions of the SEM-CL-HBPE three-model mix	115 grade 5-6 students (46.09% girls); 10-13 years old	SEM - CL-HBPE	13 lessons an educative version of CrossFit	Interviews	Qualitative: thematic analysis method	Social
Farias et al. (2015) Portugal	Analyzing the impact of SEM-IGCM on students' performance and understanding of soccer games	24 grade 5 students, comprising 16 girls and 8 boys, with a mean age of 10.3 years	SEM - IGCM	17 lessons soccer	The Game Performance Observation Instrument, Coding Association 6 Conference, The Game Understanding Test	Mixed studies: Mann-Whitney test, Wilcoxon test	Motor, cognitive
Farias et al. (2019) Portugal	To examine game performance according to the tactical structures of invasion games throughout three consecutive model-based units	26 students (10 females and 16 males); mean age of 12 years and three months	SEM - IGCM	Basketball: 20 lessons; Handball: 12 lessons; Football: 16 lessons	Game Performance Assessment Instrument (GPAI), Video observation	Quantitative: analyses of variance	Motor
Farias et al. (2022) Portugal	To explore the influence of SEM-SGA on pre-service teachers' professional content knowledge (SCK) and students' play development	60 eighth-grade students (three classes: 38 boys, 22 girls, Mage = 13.3)	SEM - SGA	10 lessons volleyball	Using audio/ video data	Quantitative: analysis of variance; Mauchly's test; Bonferroni post hoc comparisons	Motor
Fernandez-Rio &	Evaluating the	71 grade 9	SEM -	16 lessons	Open-ended	Qualitative: thematic	Social

Menendez (2017) Spain	perceptions of students and teachers in an educational kickboxing learning unit	students, with an average age of 15.4 years and a standard deviation of 0.73	TPSR	kickboxing	questions, Photovoice, teacher and external observers' diaries, semi-structured interviews	analysis method	
García-González et al. (2020) Spain	Analyzing the effects of SEM-TGFU on student initial motivations	49 students (M = 15.50, SD = 0.57)	SEM - TGFU	10 lessons volleyball	Questionnaires on basic psychological need (BPN) support and satisfaction, and intention to be physically active	Quantitative: analysis of variance	Affective
Gil-Arias et al. (2017) Spain	Assessing students' motivation levels for participating in physical activity	55 grade 9 students, comprising 27 females and 28 males, with a mean age of 15.45 years	SEM - TGFU	16 lessons volleyball	Scales: Autonomous motivation, Basic psychological needs, Enjoyment, Intention to be physically active	Quantitative: control group, MANOVA, Shapiro-Wilks test	Affective
Gil-Arias et al. (2020) Spain	Analysis of SEM-TGFU on autonomy support, sensory Knowing the effects of motivating atmosphere, fun, and perceptual ability	53 grade 9 students (16 female, 37 male); mean age 15.50	SEM - TGFU	16 lessons; handball and basketball	Physical Education Class Learning and Performance Orientation Questionnaire, Autonomy Support Coaching Strategies Questionnaire,	Quantitative: a counter-balanced crossover design	Affective

					Enjoyment and Perceived Ability Scale		
Gil-Arias, Diloy-Peña, et al. (2020) Spain	Analyzing the effects of SEM-TGFU on student motivational outcomes	53 grade 9 students (16 female, 37 male); mean age 15.50	SEM - TGFU	10 lessons volleyball	Questionnaires, focus groups	Mixed: one-way analysis of variance, analysis of variance, deductive content analysis	Affective
Gil-Arias et al. (2021) Spain	Investigating the effects of using SEM-TGFU on perceived autonomy support, perceived need satisfaction, autonomy motivation, and adaptive outcomes	292 grade 6 students (140 female, 152 male); mean age 10.41	SEM - TGFU	16 lessons basketball	Autonomy Support Coaching Strategies Questionnaire, BPNs in Sport Scale, Perceived Causality Questionnaire, Relational Goals Questionnaire, Physical Activity Class Satisfaction Questionnaire	Quantitative: analysis of variance	Affective
Gouveia et al. (2022) Portugal	Investigating the Impact of Different Pedagogical Models on Moderate-to-Vigorous Physical Activity in Physical Education Classes	17 students, 9 males, 13.7 ± 0.90 years old	SEM - TGFU	26 lessons	Objectively measure	Quantitative: Wilcoxon signed-rank test; analysis of variance	Motor
Guijarro & MacPhail. (2021) Spain	Exploring Game Performance and Game Involvement through SEM-TGFU	85 fourth- and fifth-grade students (aged 9-11)	SEM - TGFU	15 lessons basketball	Systematic observation of video recordings of students' game behavior	Quantitative: multivariate analysis of variance	Motor
Hastie & Buchanan	To assess the	United States,	SEM -	26 lessons	Independent	Qualitative: constant	Social

(2000) The United States	effectiveness of SEM-TPSR in practice and formulate a theoretical model of Empowering Sports	there are 45 boys in the sixth grade, aged 11 to 13 years old	TPSR	Xball	observations, daily debriefs, and student interviews	comparison technique	
Hastie & Curtner-Smith (2006) Australia	Analyzing the effects of implementing SEM-TGFU on both teachers and students	Australia; 29 grade 6 students, consisting of 11 boys and 18 girls, aged 11-12 years old	SEM - TGFU	22 lessons batting/fielding games	Reflective logs, critical incident reflective sheets, tactical quizzes, game design forms, and team interviews	Analytic induction technique, enumerative analysis, typological analysis, and constant comparison	Cognitive, affective
Jia (2021) China	Investigating the effects of SEM-TGFU on students' football cognitive performance and motor skills	224 students	SEM - TGFU	16-week (3 hours per week for a total of 48 hours) experimental	The ball games cognitive performance measurement (Prat et al., 2020); the benchmark of motor skills (Wang, 2018)	Quantitative: Analysis of variance	Cognitive; Motor
López-Lemus et al. (2023) Spain	Analyzing the impact of SEM-TGFU on students' enjoyment, perceived competence, intention to be physically active, skill execution, decision-making, performance, and game involvement	CG: 70 students; age = 14.43 ± 0.693; n = 32 female) EG： (SEM-TGFU): 67 students; age = 13.91 ± 0.900; n = 30 female)	SEM - TGFU	12 lessons handball	Game Performance Assessment Instrument (GPAI); the Enjoyment and perceived competence scale (ECS); the Measure of Intentionality to be Physically Active (MIFA)	Quantitative: pre-test/post-test quasi-experimental design	Cognitive; affective; Motor
Menendez &	Analyzing the impact of	12 students, with	SEM -	16 lessons	Drawings,	Qualitative: thematic	Social;

Fernandez-Rio (2017) Spain	SEM-TPSR on learners with disabilities during a contactless kickboxing learning unit	5 of them having disabilities, and their ages range from 15 to 16 years	TPSR	kickboxing	open-ended questions, discussion groups, diaries and semi-structured interviews	analysis method	Affective
Mesquita et al. (2012) Portugal	Analyze the influence of SEM-IGCM on student decision-making, skill execution, and overall performance in competition	26 grade 5 students, consisting of 17 girls and 9 boys, aged between 10 and 12 years old	SEM - IGCM	22 lessons soccer	Game Performance Assessment Instrument (GPAI), Video observation	Quantitative: Mann-Whitney test, Wilcoxon test	Motor, cognitive
Oliveros & Fernandez-Rio (2022) Spain	Investigating whether a hybrid pedagogical model could make a difference in adolescent girls' in-class physical activity levels	66 students aged 13-17 years	SEM - TGFU	12 lessons	Objectively measure	Quantitative: analysis of variance	Motor
Pan et al. (2023) China	Comparing the learning effects between SEM-TGFU and TGFU on students' motivation, enjoyment, responsibility, and game performance	90 students in the 4 classes (experimental group: 24 boys and 22 girls, Mage = 15.02 ± 0.73 years, and control group: 23 boys, and 21 girls, Mage = 14.78 ± 0.66 years)	SEM - TGFU	10 weeks with 20 PE lessons	Responsibility scale in physical education (RSPE); Learning motivation scale in physical education (LMSPE); Sport enjoyment scale in physical education (SESPE);	Quantitative: A quasi-experimental design; Analysis of covariance	Cognitive; affective; Motor

					Game performance assessment instrument (GPAI)		
Quiñonero-Martínez et al. (2023) Spain	investigating the effect of the SEM-SPRM on students' physical fitness and physical activity	76 Secondary Education students aged 12-14 (male: 32; female: 44)	SEM - SPRM	17 lessons Colpbol sport; racket games; Frisbee Ultimate	Objectively measure	Quantitative: A quasi-experimental pre-post study	Motor
Silva et al. (2022) Portugal	Investigating students' perceptions about lived learning experiences and active involvement in SEM-SGA	25 students as participants (aged be-tween16 and17 years old) who	SEM - SGA	26 lessons volleyball	Diary and interviews	Qualitative: thematic analysis method	Cognitive; effective
Stran et al. (2012) The United States	To assess pre-service teachers' perceptions of SEM-TGFU and analyze the facilitators and barriers they encountered during model implementation	22 pre-service teachers, comprising 14 males and 9 females, with an average age of 23. Additionally, there are 162 grade 5 students, aged 10-11 years old	SEM - TGFU	20 lessons Invasion games	Focus group interviews, critical incident reflections, lesson plans, and observations	Qualitative: thematic analysis method	Cognitive, affective
Wei et al. (2020) China	Analysis of the effects of SEM-TPSR on students' responsibility and exercise self-efficacy	204 students	SEM - TPSR	15-week (45 hours)	Questionnaires	Quantitative: regression analysis	Affective cognitive

Appendix B

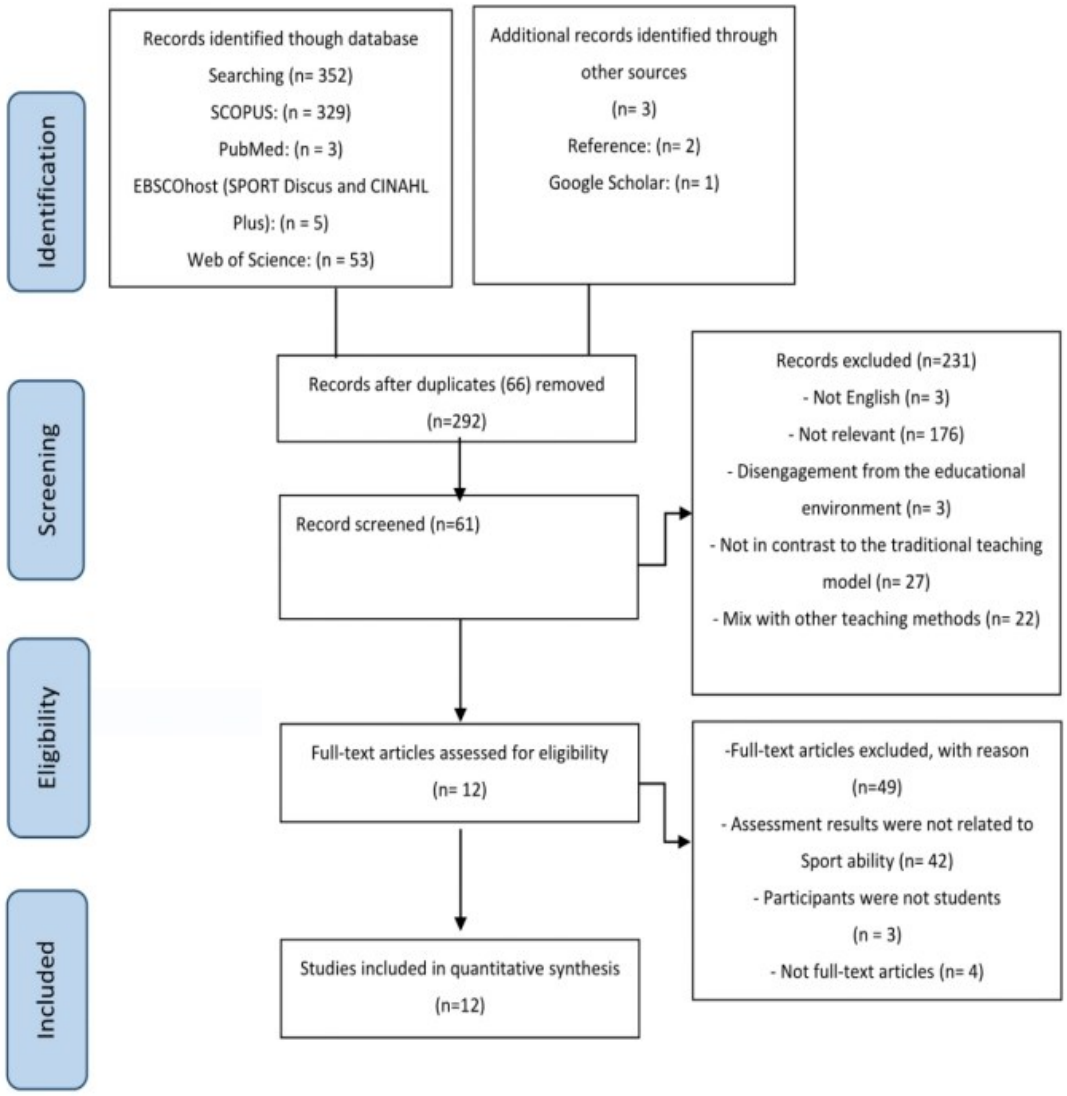


Figure 1 : Flow-Chart of Reviews' Selection Procedure

Table 1 : Characteristics of Included Studies

Author(s)/ Date	Purpose	N (sex) / Age	Country	Sport	Intervention	Comparison	Fidelity	Variables	Main Findings
Kastrena & Setiawan (2018)	To explore the influence of SEM and TT on improving students' volleyball playing skills	Middle school N=30 Male 22, Female 3	Indonesia	Volleyball	SEM 19 lessons Three times a week	TT	Unit plan and curriculum elements	volleyball playing skills	-SEM is significantly better than TT in improving students' volleyball skills
Layne & Piipari (2015)	To investigate the effect of the SEM model on basketball performance and knowledge of physical activity curriculum content among students	College N = 25, Male 22, Female 3 EG=12 male and 1 female CG=10 male and 2 female	US	basketball	SEM 28 lessons Two to three times a week One hour and 25 minutes	TT	No	Offensive game performance Game efficiency	- Students in the sports education group showed significant improvements in offensive game performance compared to the TT model. - There was a slight increase in game efficiency in the SEM group, whereas students in the traditional group experienced a decrease.
Layne (2015)	To compare the effects of SEM and DI on students' running performance	College 26 students (15 males, 11 females) SEM=15 CG=11	US	Running	SEM 42 lessons Three times a week 50min	DI	Experienced & Checklist	Running performance	- Both two teaching methods can significantly improve the jogging performance of college students. - There was no significant difference between the two teaching methods.
Li et al (2022)	To compare the effects of SEM and TT	College EG=55 (41 boys, 14 girls)	China	Volleyball	SEM 16 lessons Once a week	TT	Unit plan and curriculum	Game-like Skills (forearm	- In the SEM group, the Performance of forearm pass, overhead setting

	teaching modes on the ability and knowledge of volleyball in Chinese sports professionals.	CG=55 (39 boys, 16 girls)			90min		elements and experienced	pass, overhead setting) Game Performance (Adjustment, decision making, skill technique, and skill outcome)	skill, and Game Performance were significantly improved. - In the TI group, only the Performance of overhead setting skill and Game Performance produced a significant improvement effect. - SEM was significantly better than TI.
Nikravan et al (2019)	To compare the effects of traditional methods and sports education on Physical Fitness and Competency.	Junior EG=18 CG=18 grade two of junior school	Iran	volleyball	SEM 12 sessions 90 min	TT	No	Physical Fitness (Agility, Jump Sport, Sit-up, Speed, Endurance, Flexibility, Horizontal bar) Competency	- Comparison between groups: Exercise education was given priority over traditional methods in improving physical fitness (except jumping). - Intra-group comparison: the competency was significantly improved.
Pereira et al (2015)	To compare the influence of SEM and DI on students' track and field performance	Junior N=47 10-13 years old EG=19 (9 boys and 10 girls) CG=28 (16 boys and 12 girls)	Portugal	Track and Field	SEM 20 lessons Twice a week 45 min	DI	Unit plan and curriculum elements & Experienced & Checklist	Shot-put Triple-jump Hurdles	- The SEM group showed significant improvements in all items, but evidence of significant improvements in DI was limited to boys and taller students
Hastie et al. (2009)	To explore the effects of obstacle course	Primary school N = 48	US	Disc Lacrosse	SEM 20 classes Five times a	DI	Experienced	Aerobic fitness (PACER)	The Aerobic fitness of students in the SEM group was significantly

	education season on the aerobic fitness level of college students	(Mage=10.6), M/F=23/ 25 EG = 23 (M\F=12/11) CG = 25 (M\F=12/11)			week 40min				better than that in the DI group.
Pritchard et al (2008)	To investigate how SEM and TS would affect skill development, knowledge, and game performance for volleyball at the secondary level.	Junior EG=26 CG=21 Age=14–15	US	Volleyball	SEM 20 lessons Five times a week 50 min	TS	Unit plan and curriculum elements & Checklist	Skill (Set and Pass) Game Performance (Decision made Skill execution Adjust)	- There was no statistically significant difference between the skills, but there was a statistically significant difference between the game performance, SEM better than TS
Rocamora et al (2019)	To assess SEM and DI in Primary Education students' physical activity intensity levels, game performance, and friendship goals.	Primary school N=88 EG=47 CG=41 Mage=11.16± 0.63	Spain	Handball	SEM 15 lessons Three times a week 45min	DI	Trained& Experienced & Checklist	Game performance (Progress to goal Get open Defend opponent Prevent score)	The SEM could be considered more effective than DI to improve Primary Education students' game performance.
Wahl-Alexander & Chomentowski (2018)	To compare the effects of SEM and TI on aerobic fitness levels of college	College M=23, F=24 EG= 12 boys and 11 girls Mage=22 CG=11 boys	US	physical conditioning Sports Education course	SEM 26 lessons Bi-weekly for 75 min	TI	Unit plan and curriculum elements and experienced	Cardiovascular fitness levels	The Progressive Aerobic Cardiovascular Endurance Run (PACER) and the one-mile run were significantly improved in the SEM

	students.	and 13 girls Mage=21.25							group, and the PACER laps were significantly improved compared with the TT group.
Xu et al (2019)	To compare the effects of SEM and TSM on table tennis skills and attitudes of high school students in China.	High school N=64 Age: 16–17 years old	China	Table tennis	SEM 16 lessons Once a week 40min	TSM	Unit plan and curriculum elements and experienced	Table tennis skills (Backhand drive Forehand drive Topspin serve)	- The technical level of the students in both classes has improved significantly, and the progress of SEM students in forehand drive and serve is obviously higher than that of TSM students.
Zhang & Su (2020)	The enlightenment of the physical education model to college physical educatio.	College N=60 (20 boys, 40 girls) Mage= 20.52±0.8	China	Physical actions and running	SEM 16 lessons Once a week 90 min	TT	No	Physical quality	The SEM model was significantly better than the control group in improving the physical quality of college students.

DI: Direct instruction, TI: Traditional instruction, TS: Traditional Style, TSM: Traditional Sports Model

Appendix C

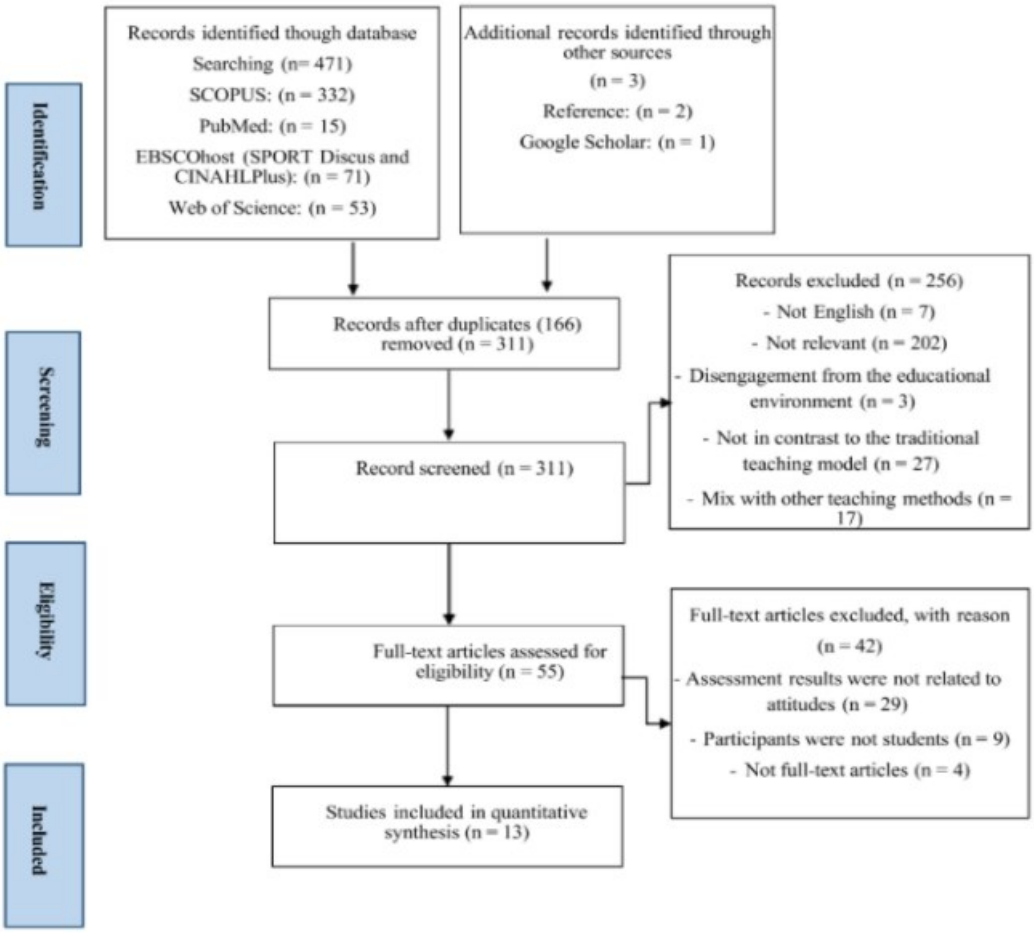


Figure 1 : PRISMA Flow

Table 1 : Detailed Search Strategy about the Systematic Review

Database	P) population (I) Sport Education Model (O) Learning Attitude	Results
Scopus	TITLE-ABS-KEY ("Sport Education Model" OR "Sport Education" OR "Sport season") TITLE-ABS-KEY ("learning attitude" OR "sports attitude" OR "cognitive" OR "cognition" OR "usefulness" OR "importance" OR "perceptions" OR "affective" OR "emotional" OR "enjoyment" OR "happiness" OR "well-being" OR "Blessedness" OR "subjective well-being") AND TITLE-ABS-KEY ("student" OR "pupil" OR "scholastic" OR "adolescent" OR "teenager")	79
EBSCohost (CENTRAL, SPORTDicus)	AB “"Sport Education Model" OR "Sport Education" OR "Sport season"” AND AB ("learning attitude" OR "sports attitude" OR "cognitive" OR "cognition" OR "usefulness" OR "importance" OR "perceptions" OR "affective" OR "emotional" OR "enjoyment" OR "happiness" OR "well-being" OR "Blessedness" OR "subjective well-being") AND AB "student" OR "pupil" OR "scholastic" OR "adolescent" OR "teenager"	46
Web of Science	((AB= ("Sport Education Model" OR "Sport Education" OR "Sport season")) AND AB= ("learning attitude" OR "sports attitude" OR "cognitive" OR "cognition" OR "usefulness" OR "importance" OR "perceptions" OR "affective" OR "emotional" OR "enjoyment" OR "happiness" OR "well-being" OR "Blessedness" OR "subjective well-being")) AND AB= ("student" OR "pupil" OR "scholastic" OR "adolescent" OR "teenager")	54
Pubmed	(("Sport Education Model" OR "Sport Education" OR "Sport season" [Title/Abstract]) AND ("learning attitude" [Title/Abstract] OR " sports attitude "[Title/Abstract] OR "cognitive"[Title/Abstract] OR "cognition "[Title/Abstract] OR "usefulness"[Title/Abstract] OR "importance"[Title/Abstract] OR "perceptions"[Title/Abstract] OR "affective"[Title/Abstract] OR "emotional"[Title/Abstract] OR "enjoyment"[Title/Abstract] OR "happiness"[Title/Abstract] OR "well-being"[Title/Abstract] OR "Blessedness"[Title/Abstract] OR "subjective well-being"[Title/Abstract])) AND ("student"[Title/Abstract] OR "pupil" [Title/Abstract] OR "scholastic" [Title/Abstract] OR "adolescent" [Title/Abstract] OR "teenager" [Title/Abstract]))	3

Table 2 : Characteristics of the Studies Examined in the Present Review

Study	Design	Participants	Interventions	Type of Teaching	Measuring Instruments and Main Outcomes
Wallhead & Ntoumanis (2004)	Pre-post Test, basketball	High school N = 51, Age:14.3 ± 0.48 yr. EG = 25 (M\F) CG = 26 (M\F)	SEM F: once a week 8 classes (60min), Same site.	TTM (TS) (England)	Intrinsic Motivation Inventory (IMI) SEM > TTM: Affective: SEM ↑, TTM ↔; Cognitive: SEM ↑, TTM ↔.
Spittle & Byrne (2009)	Pre-post Test, Hockey, soccer	Junior high N = 115, M/F, 97/ 18 Age: 13–14 yr. EG = 41 (M\F) CG = 74 (M\F)	SEM 20 classes (45min), F: Once a week Same site.	TTM (DI) (Australia)	Intrinsic Motivation Inventory (IMI) SEM > TTM: Affective: SEM ↑, TTM ↓; Cognitive: SEM ↑, TTM ↓.
Perlman (2010)	Pre-post Test, basketball	High school N= 78 EG = 40 (M\F) CG = 38 (M\F)	SEM 20 classes (60min), F: Three or four times a week，Same site	TTM (SDG) (United States)	Intrinsic Motivation Inventory (IMI) SEM >TT Affective: SEM ↑, TTM ↔.
Menickelli & Hastie (2014)	Pre-post Test, Disc Lacrosse	High school N = 40, M/F, 30/ 10 Age:15.9 ± 1.1 yr. EG = 20 (M\F) CG = 20 (M\F)	SEM 20 classes F: Five times a week Same site.	TTM (SDG) (United States)	Intrinsic Motivation Inventory (IMI) SEM > TTM (Affective and Cognitive)
Wallhead et al. (2014)	Pre-post Test, Ball course	High school M/F, 310/ 258 Age:14.74 ± 0.48 yr. EG = 281 (M\F) CG = 287 (M\F)	SEM F: Two or three times a week 25 classes (90min), Two schools.	TTM (MM) (United States)	A motivation subscale of the Academic Motivation Scale (AMS) SEM > TTM: Affective: SEM ↑, TTM ↔; Cognitive: SEM ↑, TTM ↔.
Cuevas et al. (2016)	Pre-post Test, Volleyball	High school N = 86, M/F, 49/ 37	SEM 19 classes F: Twice a week	TTM (DI) (Spain)	Intention to be Physically Active Scale (IPAS) SEM > TTM:

		Age: 15.65 ± 0.78 yr. EG = 43 (M\F) CG = 43 (M\F)	Same site		Affective: Positive affect：SEM ↔, TTM ↔; negative affect: SEM ↔, TTM ↔.
Fernández-Río et al. (2017)	Pre-post Test, Ultimate-Frisbee	High school N = 217, Age: 12-17 yr. M/F, 113/ 104	SEM 12 classes (55min), F: Twice a week， Same site.	TTM (DI) (Spain)	Intrinsic Motivation Inventory (IMI) SEM > TTM: Affective: Positive affect：SEM ↑, TTM ↔; negative affect: SEM ↓, TTM ↔.
Xu et al. (2019)	Pre-post Test, Table tennis	High school N = 64， Age: 16-17 yr. EG = 36(M\F) CG = 28 (M\F)	SEM 16 classes F: Once a week Same site	TTM (China)	The attitude questionnaire SEM > TTM: Cognitive: SEM ↑, TTM ↔; Affective: SEM ↑, TTM ↔.
Zhang & Su, (2020)	Pre-post Test, Body movements	College N = 60， M/F, 20/40 Age: 20.52 ± 0.8yr. EG = 30(M\F) CG = 30 (M\F)	SEM F: Once a week 10 classes (55min), Same site	TTM (LGM) (China)	Physical activity enjoyment scale (PACES) SEM > TTM (Affective)
Viciana et al. (2020)	Pre-post Test, Volleyball	High school N = 123, M/F, 60/ 63 Age: 14-15 yr. 3EG (M\F) 2CG (M\F)	SEM 12 classes (50min), Two different schools. F: Not mentioned	TTM (DI) (Morocco)	The Spanish version of the Sport Satisfaction Instrument (SVSSI); SEM > TTM: Affective：(negative affect) SEM ↓, TTM ↔; Cognitive: SEM ↑, TTM ↔.
Luna et al. (2020)	Pre-post Test, Polskie Ringo (a game)	primary school N = 146， M/F, 67/ 79 Age: 10.78±1.07 yr. EG = 87 (M\F) CG = 59 (M\F)	SEM 18classes (50min) F: Not mentioned Same site	TTM (DI) (Spain)	Positive and Negative Affect Scale (PANASN) SEM >TTM Affective: Positive affect：SEM ↑, TTM ↑; negative affect: SEM ↓, TTM ↔.
Chu et al. (2022)	Pre-post Test, basketball	College N = 60， M/F, 28/ 32	SEM 16classes (45min), F: Twice a week，	TTM (China)	The ARCS Learning Motivation Scale The Physical Education Affection Scale (PEAS)

		EG = 30 (M\F) CG = 30 (M\F)	Same site		SEM >TTM Affective: SEM [↑], TTM ↔, Cognitive: SEM [↑], TTM [↑] (SEM = TTM).
Iserbyt et al. (2023)	Pre-post Test, basketball	College N = 85， M/F, 70/ 15 EG = 2 classes(M\F) CG = 2 classes(M\F)	SEM 16classes (60min), Same site	TTM (Belgium)	Cognitive: SEM [↑], TTM ↔. The ALT-PE data were collected using momentary time sampling for each team by trained coders Cognitive: SEM [↑], TTM ↔.

>: The overall teaching effect has a significant advantage; [↑]: The results of pre - and post-test in the group were significantly improved; ↔: There was no significant change in the test results before and after the group; [↓]: There was no significant decrease in the test results before and after the group. N: number of participants in the experiment; M/F: male and female; EG: experimental group, CG: control group; TTM: Traditional Teaching Model; DI: Direct Instruction; LGM: Latent Growth Model; TS: Traditional Style; SDG: Skill-Drill-Game; TI: Traditional Instruction; MM: multiactivity model; F: Weekly intervention frequency. ALT-PE: Academic Learning Time Physical Education.

Appendix D

青少年体力活动问卷(PAQ-A)

姓名 ：　　出生日期 ：　　性别 ：　　班级 ：

下面的问题反映您在过去 7 天（上周）中的体力活动情况。这些活动可能让你 出汗、身体感到疲劳、呼吸急促等，如跑步、走路、爬山、舞蹈等。
请牢记：

（一） 这并不是考试，没有正确和错误答案。

（二） 请尽可能诚实、准确得回答所有问题。

1. **过去 7 天中，**你做过下列活动吗？如果做过，有多少次？多少分钟？（单项选择,请在空格内打钩） （包括体育课上、课间操、放学后）

	周一到周五，你共活动了几次					周六及周日你共活动了几次				
过去一周， 你参加过下列哪些活动	0	1-2	3-4	5-6	7 以上	0	1-2	3-4	5-6	7 以上
例如：游泳		√					√			
步行（包括上 下学）										
骑自行车（包括上下学）										
乘车上下学										
跑步										
乒乓球										
羽毛球										
足球										
篮球										
广播操										
健身操										
游泳										
舞蹈										
爬山										
跳绳										
轮滑										
跆拳道										
其他										

2. 在过去的 7 天中，你在**体育课上**活动得积极吗（游戏、跑步、跳跃、投掷）？单项选择，请在相应括号内打钩）

1） 我没上体育课（ ）
2） 很少积极活动（ ）
3） 有时很积极（ ）
4） 经常很积极（ ）
5） 一直很积极（ ）

3. 在过去的 7 天中，你在**午休期间**通常做什么（除了吃午饭以外）?（单项选择 ，请在相应括号内打钩）
   1) 坐着（包括聊天、学习、写作业、睡觉等）( )
   2) 到处站站或走走( )
   3) 跑或玩一小会儿( )
   4) 经常四处跑动或积极活动( )
   5) 大部分时间都是在跑动或剧烈运动( )

4. 在过去的 7 天中，**放学后**你有多少天积极参加体育锻炼、跳舞或其他的活动？（单项选择 ，请在相应括号内打钩）
   1) 0 天( )
   2) 1 天( )
   3) 2-3 天( )
   4) 4 天( )
   5) 5 天( )

5. 在过去的 7 天中，**有几个晚上**你积极参加了体育锻炼、跳舞或其他的活动？（单项选择 ，请在相应括号内打钩）
   1) 0 个晚上( )
   2) 1 个晚上( )
   3) 2-3 个晚上( )
   4) 4-5 个晚上( )
   5) 6-7 个晚上( )

6. 在**上周末**，你积极参加了几次体育锻炼、 跳舞或其他的活动？ （单项选择 ，请在相应括号内打钩）
   1) 0 次( )
   2) 1 次( )
   3) 2-3 次( )
   4) 4-5 次( )
   5) 6 次以上( )

7. 下面哪一句话最适合你**过去 7 天的总体情况**？请把所有 5 个选项读完再选择答案。（单项选择，请在相应括号内打钩）
   1) 我在课余时间所做的事情几乎都不需要花费体力( )
   2) 有时（上周有 1-2次）我会在课余时间参与一些体育活动（如跑步、游泳、 骑自行车、跳健美操等( )
   3) 我会经常（上周有3-4 次）地在课余时间做一些体育活动( )
   4) 我会比较经常（上周有5-6次）地在课余时间做一些体育活动( )
   5) 我会特别经常（上周有7 次或7 次以上）在课余时间做一些体育活动( )

8. 在上周的每一天，你经常进行体育活动吗（如跑步、健步走、运动锻炼、 跳舞、或者其他形式的体力活动）？（单项选择，请在相应区域内打钩）

	从来没有	一点点	一般	频繁	非常频繁
星期一					
星期二					
星期三					
星期四					
星期五					
星期六					
星期日					

9. 上周你生病了吗？或者有没有其他情况使你丌能参加正常的体育活动？（单项选择 ，请在相应括号内打钩）
   有（） 无()
   如果有， 请写出是什么事情导致你无法正常进行体育活动_____

问卷到此结束，谢谢合作！

Appendix E

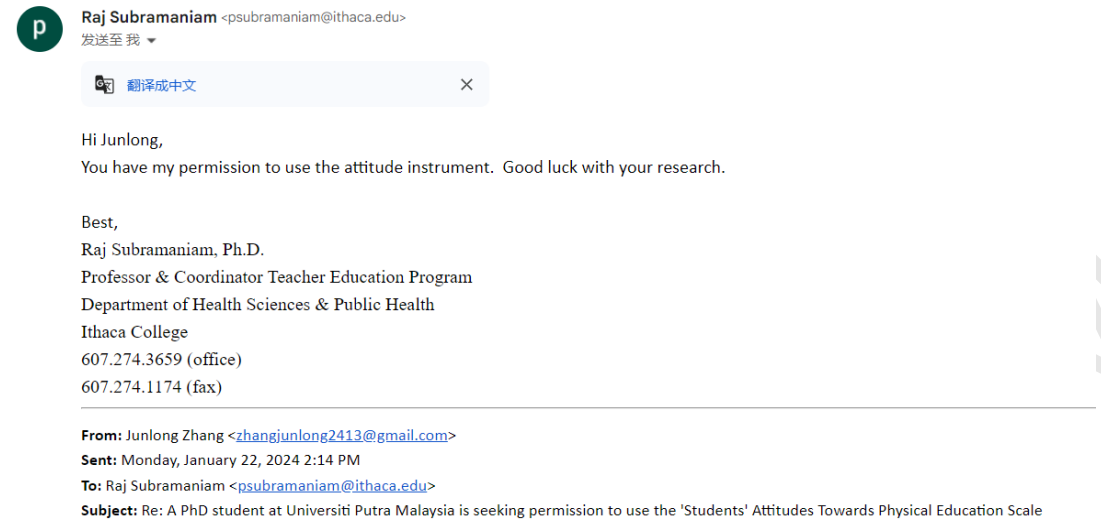
The SPEA Scale (English Version)

Question number	Problem content	Very disagree	I don't agree	same as	Compare consent	Very much agree
1	The games I learned in PE class made me feel interested in PE class.					
2	The games I learned in PE class made my study unpleasant to me.					
3	The games I learned in PE class made me feel excited about PE.					
4	My PE teacher made me feel that PE class is not important to me.					
5	I think the games I learned in PE class made PE class boring to me.					
6	I think the games I learned in PE class are useless to me.					
7	The games I learned in PE class are very important to me.					
8	My PE teacher makes me feel that PE class is very important to me.					
9	My PE teacher makes PE class fun to me.					
10	The games I learned in PE class are very useful for me.					
11	I think my PE teacher makes PE class fun.					
12	I feel like my PE teacher makes the PE class become boring to me.					
13	I think the games I learned in PE class are very valuable to me.					
14	The games I learned in PE class are not important to me.					
15	My PE teacher makes study become unpleasant in PE class.					
16	My PE teacher makes PE class very useful for me.					
17	I think my PE teacher makes PE class valuable.					

18	I don't think my PE teacher makes PE class useless to me.					
19	My PE teacher makes me feel excited about the PE class.					
20	I think the games I learned in PE class make learning interesting.					

The SPEA Scale (Chinese version)

问题编号	问题内容	非常不同意	不太同意	一般	比较同意	非常同意
1	在体育课上学到的游戏让我对体育课感到有趣。					
2	在体育课上学到的游戏让学习对我来说变得不愉快。					
3	在体育课上学到的游戏让我对体育感到兴奋。					
4	我的体育老师让我觉得体育课对我来说不重要。					
5	我觉得在体育课上学到的游戏让体育课对我来说变得无聊。					
6	我觉得在体育课上学到的游戏对我来说没有用。					
7	在体育课上学到的游戏对我来说很重要。					
8	我的体育老师让我觉得体育课对我来说很重要。					
9	我的体育老师让体育课对我来说很有趣。					
10	在体育课上学到的游戏对我来说很有用。					
11	我觉得我的体育老师让体育课变得有趣。					
12	我觉得我的体育老师让体育课对我来说变得无聊。					
13	我觉得在体育课上学到的游戏对我来说很有价值。					
14	在体育课上学到的游戏对我来说不重要。					
15	我的体育老师让学习在体育课上变得不愉快。					
16	我的体育老师让体育课对我来说很有用。					
17	我觉得我的体育老师让体育课变得有价值。					
18	我觉得我的体育老师让体育课对我来说没有用。					
19	我的体育老师让我对体育课感到兴奋。					
20	我觉得在体育课上学到的游戏让学习变得有趣。					



Raj Subramaniam <psubramaniam@ithaca.edu>
发送至 我

翻译成中文

Hi Junlong,
You have my permission to use the attitude instrument. Good luck with your research.

Best,
Raj Subramaniam, Ph.D.
Professor & Coordinator Teacher Education Program
Department of Health Sciences & Public Health
Ithaca College
607.274.3659 (office)
607.274.1174 (fax)

From: Junlong Zhang <zhangjunlong2413@gmail.com>
Sent: Monday, January 22, 2024 2:14 PM
To: Raj Subramaniam <psubramaniam@ithaca.edu>
Subject: Re: A PhD student at Universiti Putra Malaysia is seeking permission to use the 'Students' Attitudes Towards Physical Education Scale

Figure 1 : Permission of SPEA Scale

Appendix F

The Test Methods and Instruments of Physical Fitness:

a) Body composition
   - ➢ Equipment needed: Height and weight measurement device
   - ➢ Testing requirements: Height and Weight: Subjects remove their shoes and wear light clothing. They stand upright on the measuring device, take a deep breath, hold it, stand straight, and look forward. The tester records the subject's height and weight measurements, with height recorded in centimeters (cm) and weight in kilograms (kg), accurate to one decimal place.

b) Flexibility: Sit and reach test
   - ➢ Equipment needed: Sit and reach box testing
   - ➢ Requirements: The participant sits on a mat, with feet against the sit and reach the box, knees extended. Slowly bending forward, the participant reaches forward with the fingertips of both hands, pushing the slider of the equipment until maximum range without sudden force. Knees must remain extended without bending. Retest if there are significant technical quality issues. The measurement unit is in centimeters, accurate to one decimal place.

c) Agility: Arrowhead agility test
   - ➢ Facility and equipment needed: Football field, six cones/markers, one tape measure, electronic timing gates or stopwatch
   - ➢ Dress lightly and ensure a thorough warm-up to ensure the safety of the testing environment.
     Left-side Test: The athlete starts from the starting point and runs forward to point A, then makes a right turn, and runs to point D. After reaching point D, the athlete makes a left turn and runs to point B. Upon reaching point B, the athlete makes another left turn and runs back to the starting point, completing the test. Right-side Test: The athlete starts from the starting point and runs forward to point A, then makes a left turn and runs to point C. After reaching point C, the athlete turns right and runs to point B. Upon reaching point B, the athlete turns right and returns to the starting point, completing the test. Each participant must complete both left and right tests with an interval of at least 3 minutes between each test. Record the times for both tests separately (marking left and right). Each athlete performs the test twice, and the best time is recorded. The unit is seconds, accurate to two decimal places.

d) Speed: 30-meter sprint

- ➢ Equipment needed: football field, tape measure, electronic timing gates or stopwatch, cones or poles (to mark the start, midpoint, and finish lines).
- ➢ Testing requirements: Dress lightly and ensure a thorough warm-up to ensure the safety of the testing environment.
  Start from a standing position and sprint the 30-meter distance as fast as possible. Each athlete performs the test twice, and the best time is recorded. Record the times at 5, 10, and 20 meters as reference points. The unit of measurement is seconds, accurate to two decimal places.

e) Upper body strength: pull-up
- ➢ Equipment needed: pull-up bar
- ➢ Testing requirements: The participant grips the bar with palms facing forward, hands slightly wider than shoulder-width apart. During the pulling-up process, slight body swinging is allowed. The participant pulls up until the chin clears the bar. Count the number of completed pull-ups according to the requirements. Pull-ups with significant technical quality issues are not counted.

f) Lower body explosiveness: vertical touch jump.
- ➢ Equipment needed and testing requirements: Find a wall with a height greater than 3 meters and mark it with a ruler. The participant raises one arm overhead, extends the arm, and measures the height from the fingertips to the ground, recording it. The participant stands vertically against the wall, arms hanging down, then jumps vertically upwards (without a run-up), reaching up with one hand as high as possible against the wall. Record the height touched. The participant has three attempts, and the highest value is recorded. The measurement unit is in centimeters. Vertical jump height = height touched - height from fingertips to ground. Any alternative methods used should be noted.

g) Cardiovascular fitness: Yo-Yo test
- ➢ Equipment needed: Football field, tape measure, cones (to mark start and finish lines, with intervals of 20m), Yo-Yo test software, and player.
- ➢ Testing requirements: Ensure thorough warm-up and safety of the testing environment.
  Group athletes and begin the test. Select appropriate mode: 13-15 years: IR1; 16-22 years: IR2.
  Athletes perform shuttle runs with increasing speed between two cones placed 20 meters apart, with intermittent beeps from a recording device controlling the rhythm (confirm correct test mode before starting). Athletes must complete each shuttle run between two beeps. If unable to complete, they will be reminded once; if still unable, the test ends after the second reminder. The last completed shuttle run before the end of the test is recorded, and the total distance covered is measured in meters.

Appendix G

Socio-Demographic Information

➢ **Basic information**

Name：_____ Date of birth: _____
Height(m): _____ Weight (kg): _____
Grade: _____ Schools: _____

➢ **Subjective Characteristics**

➢Family income level (yuan)

_____ ＜3000
_____ 3001-6000
_____ 6001-9000
_____ 9001-12000
_____ 12001-15000
_____ ＞15001

➢Family Cultural background:

_____City
_____Village
_____Others, please specify:

➢Chronic health problems or disabilities: _____Yes _____No

➢Self-rated health: _____Yes _____No

➢Health behaviour

_____Exercise for at least 1 hour every day
_____Exercise for less than 1 hour every day
_____No regular exercise habits

Table 1 : Assessment of Instrument Validity

Instruments	Relevance				Clarity			
	Total Agreement	Number of experts	I-CVI	KAPPA	Total Agreement	Number of experts	I-CVI	KAPPA
Socio-demographic Information	5	6	0.833-1.000	0.816-1.000	6	6	1	1
Physical fitness test	5	6	0.833-1.000	0.816-1.000	6	6	1	1
Students' Attitudes towards Physical Education Scale (SPEA)	6	6	1	1	6	6	1	1

Appendix H

Protocol

Validation of Experts for Intervention

1. Basic Information of expert

Expert Information	
Work Unit:	
Full name:	
Title:	
Research Area:	

Dear expert,

My name is Zhang Junlong, and I come from the University Putra Malaysia. The research topic is "Effect of Football's Sport Education Model as Intervention on Physical Fitness and Learning Attitude Among Junior High School Students in Hebei, China." You have a wealth of theoretical knowledge and practical experience. This protocol wants you to validate whether this 12-week teaching intervention is reasonable. Your suggestions will only be used to improve the experimental design and not for any other purpose. All information you give will be kept confidential.

Thank you so much for being so helpful, and best wishes to you.

Universiti Putra Malaysia
Supervisor committee:
Prof. Dr. Soh Kim Geok
Dr. Mohd Ashraff Mohd Anuar

2. Comments on the effectiveness of intervention programs of 12-week group and standard teaching group

2.1 Expert comment questionnaire

This questionnaire adopts a 4-Score (Zamanzadeh et al., 2015; Davis, 1992), with 1-4 (representing increasing support (1 = not relevant/clarity, 2 = somewhat relevant/clarity, 3 = quite relevant/clarity, 4 = highly relevant/clarity). Please score the brisk Sport Education Model intervention. If there are other contents, you can add them and indicate the name. After reading the questionnaire, please rate the following aspects and tick "on the corresponding rating scale. (See Table 1 for an Evaluation of the effectiveness of the intervention program).

Table 1 : Evaluation of the Effectiveness of the 12-week Intervention Program

Type	Relevant				Clarity				Improvement suggestions
	4	3	2	1	4	3	2	1	
Overall evaluation of the program									
Duration									
Frequency									
Intensity									
Time									
Warm-up									
Basic content									
Cool down									
Rest time									
Risk									

2.2 Intervention Plan-Same Characteristics

Table 2 : Intervention Plan-Same Characteristics

Content	Sport Education Model Group and Standard Teaching Group	Improvement suggestions
Teacher qualification	Same qualification and titles	
	Both have one teacher	
	Same majors: physical education Same experience: Minimum more than 10 years	
Student qualifications	Similar age groups：13-15 Same background：ordinary junior high school No students have received this course prior to the study	

2.3 Comparison of teaching characteristics between The Sport Education Model and Standard Teaching

The six characteristics of The Sport Education Model include: season, affiliation, formal competition, culminating event, Record-keeping, and Festivity (Siedentop, 2002；Bessa et al., 2019).

Table 3 : Main Characteristics of Both Instructional Approaches

Sport Education Model	Standard teaching
Season: organize the entire teaching unit in the form of a football league	Teach in the usual instructional format
Affiliation: each student was randomly assigned to a fixed team and played a role: player, coach, conditioning coach, captain, referee, and equipment manager	Students have no fixed teams and do not play roles, just practice individual according to what the teacher teaches.
Formal competition: each fixed team plays in the knockout round.	There is no formal competition, but there is game play, and the time spent usually depends on the level of success in practice.
Culminating event: formally announce the	Do not formally announce match results and

results and points of each formal competition and organize the culminating event (final).	do not define finals.
Record-keeping: record and formally announce the results of each team, including capturing photographs and videos.	Game scores were not kept officially, although the winners of individual games may have broadcast them informally.
Festivity: review the entire sports season; Award the winning teams and individuals Course Summary	There's no festivity.

2.4 Teaching plan of the Sport Education Model and Standard Teaching

2.4.1 Intervention Prescription

Table 4 Intervention Prescription of the Sport Education Model and Standard Teaching

Week	Frequency	Type of exercise	Time				Intensity	Improvement suggestions
			Warm-up (min)	Basic content (min)	Cool-down (min)	Duration (min)	HRmax	
1	3 / week	SEM	10	10min-rest for 5min-10min	5	40	HRmax (50%-55%)	
		NT		10min-rest for 5min-10min			HRmax (50%-55%)	
2-5	3 / week	SEM	10	10min-rest for 5min-10min	5	40	HRmax (60%-70%)	
		NT		10min-rest for 5min-10min			HRmax (60%-70%)	
6-7	3 / week	SEM	10	10min-rest for 5min-10min	5	40	HRmax (70%-80%)	
		NT		10min-rest for 5min-10min			HRmax (70%-80%)	
8-9	3 / week	SEM	10	10min-rest for 5min-10min	5	40	HRmax (75%-85%)	
		NT		10min-rest for 5min-10min			HRmax (75%-85%)	
10-11	3 / week	SEM	10	10min-rest for 5min-10min	5	40	HRmax (80%-90%)	
		NT		10min-rest for 5min-10min			HRmax (80%-90%)	
12	3 / week	SEM	10	10min-rest for 5min-10min	5	40	HRmax (60%-70%)	
		NT		10min-rest for 5min-10min			HRmax (60%-70%)	

SEM: Sport Education Model; ST: Standard Teaching

According to the teaching framework characteristics of the Sport Education Model, the twelfth week is designated as the Festivity for the Culminating Event (Siedentop, 2002). Its main content should include: reviewing the entire sports season, awarding the winning teams and individuals, and course summary (Bessa et al., 2019; Pan, 2023). Because the activities during the Festivity are not aimed at enhancing physical fitness, it results in a relatively low intensity for the 12-week period. Its purpose is to promote students' learning attitudes (affective and cognitive) and to help them develop as competent, literate, and enthusiastic sportspersons (Bessa et al., 2019).

2.4.2 Details of teaching program

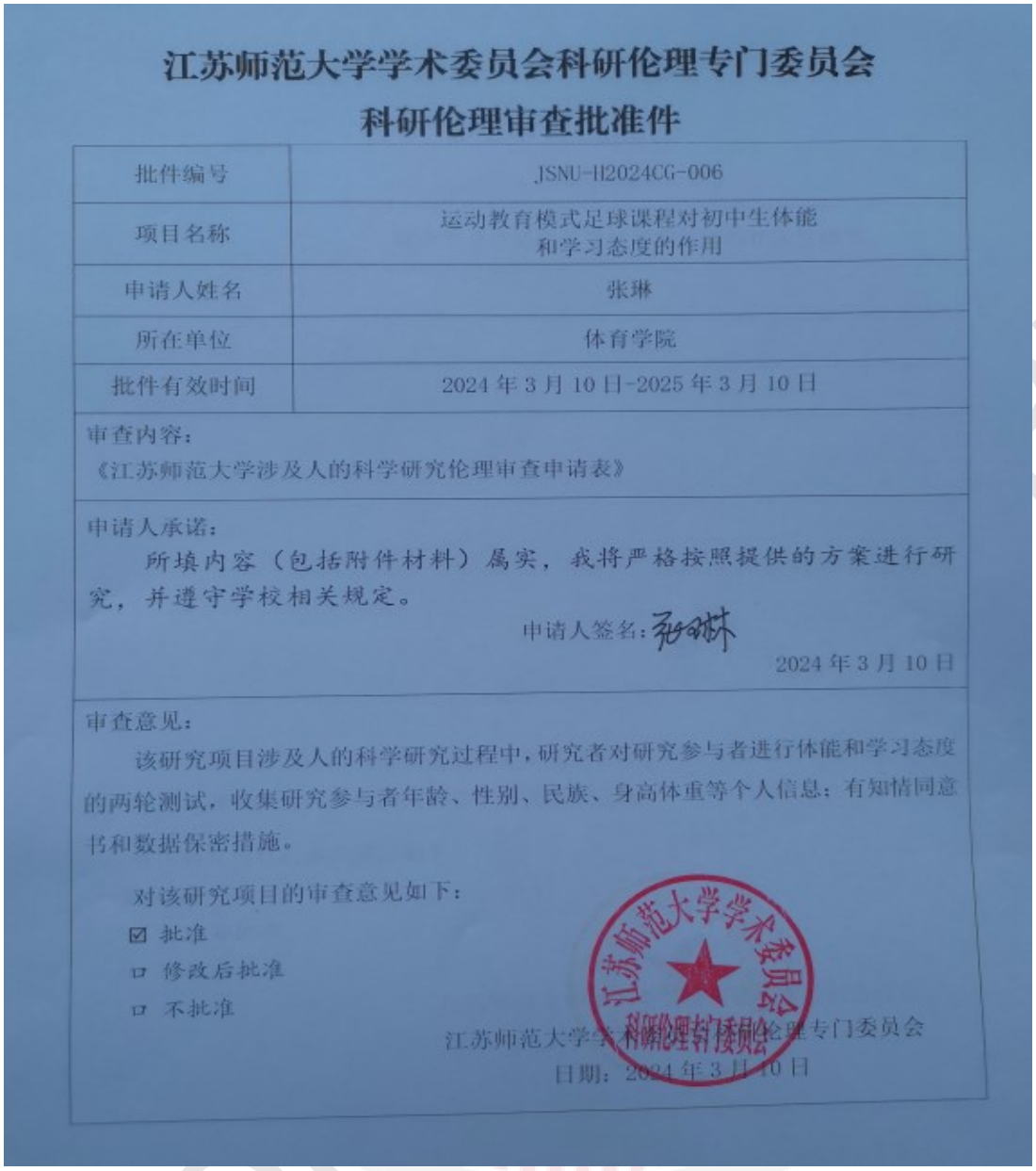
The Teaching experiment period of Sport Education Model and Standard Teaching was 12 weeks (Pan, 2023; Oliveros & Fernandez-Rio, 2022; Lopez-Lemus, 2023), 40 minutes per lesson which includes warm-up (10-min), basic content (20min), rest of basic content (5min), and cool-down (5min) (Pan, 2023; Qu JiaLiang, 2022). The most significant difference between the two teaching methods lies in delivering basic content. The Sport Education Model includes Teacher-directed learning (10 minutes) and Team practice (10 minutes); Standard Teaching includes Teacher-directed learning (10 minutes) and Individual practice (10 minutes).

Table 5 : Details of Teaching Program of the Sport Education Model and Standard Teaching

<table>
<tr><th colspan="5">Sport Education Model</th><th colspan="4">Standard Teaching</th></tr>
<tr><td colspan="9">Warm-up (10-min)</td></tr>
<tr><td colspan="9">➢ Light Jogging (2-min)
➢ Dynamic Stretching (3-min)
(a) High knees (2 sets, 20 sec/ set, 20 sec for rest between sets)
(b) Butt kicks (2 sets, 20 sec/ set, 20 sec for rest between sets)
(c) Side shuffles (2 sets, 15 reps/ set, 20 sec for rest between sets)
➢ Agility Drills with Football Ball (4-min):
(a) Dribbling drills (2 sets, 40 sec/ set, 40 sec for rest between sets)
(b) Passing drills (2 sets, 40 sec/ set, 40 sec for rest between sets)
➢ Bodyweight squats (1-min) (1 set, 20 reps, 20 sec for rest)
➢ Ball Control Drills (1-min) (1 set, 40 sec, 20 sec for rest)</td></tr>
<tr><td colspan="9">Basic content (20min)</td></tr>
<tr><th rowspan="2">Week</th><th rowspan="2">Content</th><th colspan="2">Time</th><th rowspan="2">Week</th><th rowspan="2">Content</th><th colspan="2">Time</th></tr>
<tr><th>Teacher-directed learning</th><th>Team practice</th><th>Teacher-directed learning</th><th>Individual practice</th></tr>
<tr><td>1</td><td>The teacher introduces the purpose of the study, expectations and establishes affiliations and roles within 6 teams (6-7 students per team).</td><td>20-min for 1st lesson of 1st week</td><td>_</td><td>1</td><td>The teacher introduced the purpose of the study, expectations</td><td>20-min for 1st lesson of 1st week</td><td>_</td></tr>
<tr><td rowspan="9">1-9</td><td>Stationary inside foot pass</td><td rowspan="9">10-min/ lesson,
3 lesson/ week
(Learn one technique / week)</td><td rowspan="9">4-min / set,
2 sets / lesson,
3 lesson / week
1-min rest between sets
(Learn one technique / week)</td><td rowspan="9">1-9</td><td>Stationary inside foot pass</td><td rowspan="9">10-min/ lesson,
3 lesson/ week
(Learn one technique / week)</td><td rowspan="9">4-min / set,
2 sets / lesson,
3 lesson / week
1-min rest between sets
(Learn one technique / week)</td></tr>
<tr><td>Throwing in from the sideline</td><td>Throwing in from the sideline</td></tr>
<tr><td>Receiving and controlling airborne balls while moving</td><td>Receiving and controlling airborne balls while moving</td></tr>
<tr><td>Dribbling and pulling the ball with different parts of the foot.</td><td>Dribbling and pulling the ball with different parts of the foot.</td></tr>
<tr><td>Front cutting and front steal</td><td>Front cutting and front steal</td></tr>
<tr><td>Side steal and side interception</td><td>Side steal and side interception</td></tr>
<tr><td>Covering (Zone Defense)</td><td>Covering (Zone Defense)</td></tr>
<tr><td>Double Teaming</td><td>Double Teaming</td></tr>
<tr><td>Quick Counterattack</td><td>Quick Counterattack</td></tr>
<tr><td>10</td><td>Formal competition: team 5v5 matches (See notes for competition method); Record-keeping</td><td colspan="2">10-min (first-half)
10-min (second half)
3 matches / week</td><td>10</td><td>Teacher randomly divides the students into groups for scrimmage</td><td colspan="2">10-min (first-half)
10-min (second half)
3 matches / week</td></tr>
<tr><td>11</td><td>Culminating event: team 5v5 matches (See notes for competition method); Record-keeping</td><td colspan="2">10-min (first-half)
10-min (second half)
3 matches / week</td><td>11</td><td>Teacher randomly divides the students into groups for scrimmage</td><td colspan="2">10-min (first-half)
10-min (second half)
3 matches / week</td></tr>
<tr><td>12</td><td>Festivity: review the entire sports</td><td colspan="2">First lesson of this week: Team skill</td><td>12</td><td>Review and course summary</td><td colspan="2">First and second lessons of this week:</td></tr>
</table>

	season; award the winning teams and individuals course summary	demonstration (3-min/ team / 6 teams) (2 minutes for replenishing water) Second lesson of this week: Award ceremony and acceptance speech (Team: 1.5-min/ team/ 6 teams; Outstanding individual: 1-min/ student/ 10 students) Third lesson this week: Summary of captain (1.5-min/ captain/ 6 captains); Teachers summarize (10-min)			Personal skill display (1- minute each student/ 20-min/ lesson) Third lesson this week: Teachers summarize (20- min)
Rest of basic content (5min)					
➢ Seated Rest (2-min) ➢ Hydration Break (1-min) (About 200 ml) ➢ Breathing- in out (1-min) (10 reps) ➢ Light Stretching (1-min) (10 reps)					
Cool-down (5min)					
➢Light Jog (1-min) ➢ Static Stretching (2.5-min) Hamstring stretch (20 sec for each leg) Quadriceps stretch (20 sec for each leg) Calf stretch (20 sec for each leg) Groin stretch (30 sec) ➢Deep Breathing (30 sec) ➢Reflect and Hydrate (1-min)					

Noted: The teaching content of the basic parts refers to the football teaching content and requirements in the "Compulsory Education Physical Education and Health Curriculum Standards" issued by the Chinese government in 2022 (pages 35-37). In the SEM group, the Formal competition and Culminating event are divided into two tiers: Division A and Division B leagues, with elimination rules. During the Formal competition phase, taking six groups per class as an example, teams for the first three matches are determined by draw. Winning teams advance to Division A, while losing teams move to Division B. In the subsequent culminating event, teams for the first and second matches are grouped by draw. After the second match, the third-place team is determined, followed by the contest for the championship and runner-up.



江苏师范大学学术委员会科研伦理专门委员会

科研伦理审查批准件

批件编号	JSNU-H2024CG-006
项目名称	运动教育模式足球课程对初中生体能和学习态度的作用
申请人姓名	张琳
所在单位	体育学院
批件有效时间	2024 年 3 月 10 日-2025 年 3 月 10 日

审查内容：

《江苏师范大学涉及人的科学研究伦理审查申请表》

申请人承诺：

所填内容（包括附件材料）属实，我将严格按照提供的方案进行研究，并遵守学校相关规定。

申请人签名：张琳

2024 年 3 月 10 日

审查意见：

该研究项目涉及人的科学研究过程中，研究者对研究参与者进行体能和学习态度的两轮测试，收集研究参与者年龄、性别、民族、身高体重等个人信息；有知情同意书和数据保密措施。

对该研究项目的审查意见如下：

☑ 批准

□ 修改后批准

□ 不批准

江苏师范大学学术委员会科研伦理专门委员会

日期：2024 年 3 月 10 日

Figure 1 : Ethical Approval

Appendix I

The Results of Content Validity for Intervention

The calculation method and importance of project Content Validity Index (I-CVI) and Kappa correction coefficient. I-CVI evaluates each item through a four-point scale, and experts score whether it is relevant. The higher the score, the higher the degree of expert consensus. The Kappa correction coefficient is used to determine the magnitude of the degree of agreement, and 0.78 or more is considered an acceptable level of agreement.

Table 1 : Relevancy and Agreement of the intervention

Type	Total Agreement	Number of experts	Relevant		Total Agreement	Number of experts	Clarity		Improvement suggestions
			I-CVI	KAPPA			I-CVI	KAPPA	
Overall evaluation of the program	6	6	1.000	1.000	6	6	1.000	1.000	
Duration	6	6	1.000	1.000	6	6	1.000	1.000	
Frequency	6	6	1.000	1.000	6	6	1.000	1.000	
Intensity	5	6	0.833	0.816	5	6	0.833	0.816	
Time	6	6	1.000	1.000	6	6	1.000	1.000	Both groups have the same class time
Warm-up	5	6	0.833	0.816	5	6	0.833	0.816	
Basic content	6	6	1.000	1.000	5	6	0.833	0.816	Model fidelity
Cool down	6	6	1.000	1.000	6	6	1.000	1.000	
Rest time	6	6	1.000	1.000	6	6	1.000	1.000	
Risk	5	6	0.833	0.816	5	6	0.833	0.816	

Appendix J

Table II. Instructional Checklist (Hastie et al. 2013).

1. Groups of students go to a designated home area and begin warming up with that group.
2. Students warm-up as a whole class under the direction of the teacher.
3. Students practice together with their group/team under the direction of a peer leader.
4. Students practice individually or in small groups under the direction of the teacher.
5. Students remain a part of easily identifiable groups throughout the lesson and throughout different tasks.
6. Student grouping throughout the lesson is variable across tasks.
7. Performance records are kept by students.
8. Students perform specialized tasks within their group/team.
9. Student performance scores count toward a formal and public scoring system.
10. Student performance scores are not recorded or are recorded in private.

Note. Items 1, 3, 5, 7 and 8 suggest a SE season, whilst items 2, 4, 6, and 10 are features of the TT.

Instruments

Appendix K

Table 1 : Normality Test Results of Sociodemographic Information

	Groups	Levene Statistic	df	P-value
Family income level	Control group	0.910	42	0.003
	Experimental group	0.910	42	0.003
Educational background	Control group	0.636	42	0.001
	Experimental group	0.634	42	0.001
Health 1	Control group	—	42	—
	Experimental group	—	42	—
Health 2	Control group	—	42	—
	Experimental group	—	42	—
Health behavior	Control group	0.636	42	0.001
	Experimental group	0.636	42	0.001
Height	Control group	0.963	42	0.194
	Experimental group	0.968	42	0.287
Weight	Control group	0.888	42	0.001
	Experimental group	0.938	42	0.025

Table 2 : Normality Test Results of Dependent Variables

Variables	Times	Groups	Levene Statistic	df	P-value
BMI-for-age	Pre-test	Control group	0.875	42	0.001
		Experimental group	0.917	42	0.005
	Post-test1	Control group	0.881	42	0.001
		Experimental group	0.912	42	0.003
	Post-test2	Control group	0.872	42	0.001
		Experimental group	0.914	42	0.004
Sit and reach	Pre-test	Control group	0.936	42	0.020
		Experimental group	0.968	42	0.284
	Post-test1	Control group	0.957	42	0.113
		Experimental group	0.971	42	0.352
	Post-test2	Control group	0.957	42	0.119
		Experimental group	0.957	42	0.113
Arrowhead agility test	Pre-test	Control group	0.983	42	0.790
		Experimental group	0.947	42	0.052
	Post-test1	Control group	0.967	42	0.264
		Experimental group	0.960	42	0.146
	Post-test2	Control group	0.940	42	0.028
		Experimental group	0.969	42	0.301
30-meter sprint	Pre-test	Control group	0.872	42	0.001
		Experimental group	0.956	42	0.106
	Post-test1	Control group	0.925	42	0.009
		Experimental group	0.932	42	0.015
	Post-test2	Control group	0.940	42	0.029
		Experimental group	0.940	42	0.028
Pull-up	Pre-test	Control group	0.493	42	0.001
		Experimental group	0.515	42	0.001
	Post-test1	Control group	0.575	42	0.001
		Experimental group	0.645	42	0.001
	Post-test2	Control group	0.711	42	0.001

Vertical jump	Pre-test	Experimental group	0.752	42	0.001
		Control group	0.923	42	0.008
	Post-test1	Experimental group	0.952	42	0.078
		Control group	0.950	42	0.065
	Post-test2	Experimental group	0.952	42	0.078
		Control group	0.981	42	0.693
		Experimental group	0.968	42	0.281
YOYO test	Pre-test	Control group	0.851	42	0.001
		Experimental group	0.908	42	0.002
	Post-test1	Control group	0.700	42	0.001
		Experimental group	0.899	42	0.001
	Post-test2	Control group	0.874	42	0.001
		Experimental group	0.930	42	0.013
Affective Dimension Score	Pre-test	Control group	0.942	42	0.035
		Experimental group	0.964	42	0.209
	Post-test1	Control group	0.964	42	0.209
		Experimental group	0.968	42	0.290
	Post-test2	Control group	0.973	42	0.428
		Experimental group	0.954	42	0.091
Cognitive Dimension Score	Pre-test	Control group	0.957	42	0.115
		Experimental group	0.966	42	0.234
	Post-test1	Control group	0.975	42	0.494
		Experimental group	0.977	42	0.530
	Post-test2	Control group	0.950	42	0.066
		Experimental group	0.950	42	0.066
Learning Attitude Score	Pre-test	Control group	0.948	42	0.056
		Experimental group	0.977	42	0.545
	Post-test1	Control group	0.974	42	0.456
		Experimental group	0.959	42	0.131
	Post-test2	Control group	0.970	42	0.318
		Experimental group	0.970	42	0.333

Table 3 : Tests for Homogeneity

Variables	Times	Levene Statistic	df 1	df 2	P-value
Height	Pre-test	0.409	1	82	0.524
Cognitive Dimension Score	Post-test1	12.830	1	82	0.001
	Post-test2	0.919	1	82	0.341
	Pre-test	1.537	1	82	0.219
Learning Attitude Score	Post-test1	7.724	1	82	0.007
	Post-test2	3.029	1	82	0.086
	Pre-test	0.380	1	82	0.539

Table 4 : Test of One-Way ANOVA

Variables	Sum of squares	df	Mean square	F	P-value
Height	0.003	1	0.003	0.500	0.482

Table 5 : Test of Nonparametric

Variables	Kruskal-Wallis H	df 1	df 2	Asymptotic significance
Family income level	0.048	1	82	0.826
Educational background	0.047	1	82	0.828
Health 1	0.001	1	82	1.000
Health 2	0.001	1	82	1.000
Health behavior	0.001	1	82	1.000
Weight	0.090	1	82	0.764
BMI-for-age	0.009	1	82	0.925
Sit and reach	0.085	1	82	0.771
Arrowhead agility test	1.201	1	82	0.273
30-meter sprint	0.025	1	82	0.876
Pull-up	0.685	1	82	0.408
Vertical jump	1.700	1	82	0.192
YOYO test	2.097	1	82	0.148
Affective Dimension Score	0.022	1	82	0.882
Cognitive Dimension Score	0.471	1	82	0.492
Learning Attitude Score	0.234	1	82	0.629

BIODATA OF STUDENT

Junlong Zhang was born on June 16, 1991, in Hebei, China. He completed his primary and secondary education in 2004 and 2007, respectively. He began training at a football club during his middle school years, as he was a football enthusiast. The period of professional training from ages 12 to 15 remains a memorable time for him. In 2007, he was admitted to Funing No. 1 High School due to his excellent academic and sports test scores. After graduating from high school in 2010, he pursued a Bachelor's degree in Physical Education at Shangqiu Normal University in Henan Province. After earning his Bachelor's degree in 2014, he continued with graduate studies. Upon completing his Master's degree, he served as a physical education teacher at a school, where he led the school' s football team to achieve outstanding results multiple times. Currently, he is pursuing a PhD in the Department of Sports at Universiti Putra Malaysia.

LIST OF PUBLICATIONS

Publication-Accepted/Published

Zhang, J., Soh, K. G., Anuar, M. A. B. M., Lin, X., & Nasiruddin, N. J. M. (2023). The Effect of the Sport Education Model on Sports Ability: A Critical Systematic Review. Revista de Psicología del Deporte (Journal of Sport Psychology), 32(3), 339-354.

Zhang, J., Xiao, W., Soh, K. G., Yao, G., Anuar, M. A. B. M., Bai, X., & Bao, L. (2024). The effect of the Sport Education Model in physical education on student learning attitude: a systematic review. BMC Public Health, 24(1), 949.

Junlong Zhang, Kim Geok Soh, Xiaorong Bai, Mohd Ashraff Mohd Anuar, Wensheng Xiao.Optimizing learning outcomes in Physical Education: A comprehensive systematic review of hybrid pedagogical models integrated with the Sport Education Model. (Formal Acceptance by PLOS ONE; PONE-D-24-20221R1)

Bao, L., NASIRUDDIN, N. J. M., SOH, K. G., Xie, H., Lin, X., & Zhang, J. (2023). Effect of Teaching Interventions on Metacognitive Skills in Physical Education. Int. J. Sport Psychol, 54, 509-538.

Lin, X., Samsudin, S., Geok, S. K., Nasiruddin, N. J. M., Bao, L., & Zhang, J. (2023). Psychometric Properties and Measurement Invariance of the 3× 2 Achievement Goals Questionnaire for Exercise (3× 2 AGQ-E) among Chinese Mainland adolescents. Revista de Psicología del Deporte (Journal of Sport Psychology), 32(2), 351-361.

Bao, L., Soh, K. G., Mohd Nasiruddin, N. J., Xie, H., & Zhang, J. (2024). Unveiling the Impact of Metacognition on Academic Achievement in Physical Education and Activity Settings: A Comprehensive Systematic Review and Meta-Analysis of Qualitative Insights. Psychology Research and Behavior Management, 973-987.

Publications -Under Review

Zhang J.L., Soh K.G., Mohd Anuar M.A., Xiao W.S. Bai X.R. (2024) From Mind to Body: An Umbrella Review of Pedagogical Models from an Embodied Cognition Perspective. PLOS ONE Since 12 April, 2024

Zhang J.L., Soh K.G., Mohd Anuar M.A., Xiao W.S. Bai X.R. (2024) Effects of Sport Education Model on Students' Enjoyment and Fun: A Systematic Review and Meta-Analysis. BMC Public Health Since 05 Sep 2024