*Corresponding Author's Email: 221123150143@zjut.edu.cn
Proceedings of the International Education Conference
Vol. 3, Issue. 1, 2025, pp. 1-22
DOI: https://doi.org/10.33422/ieconf.v3i1.1656

ISSN: 3030-0207 online

Empirical Research on the Development of a Fahp-Based Evaluation System and Assessment Tool for Children's Time Management Skills

Sija Wu
School of Design and Architecture, Zhejiang University of Technology, China

Abstract

Existing time management assessment systems or tools mainly rely on self-report questionnaires, which lack objectivity and fail to fully capture children's actual behaviours and developmental levels in everyday contexts. To address the absence of a scientific, task-based evaluation framework for children in the preschool-to-primary transition stage, this study constructs a three-tier indicator system for time management ability using the Fuzzy Analytic Hierarchy Process (FAHP) and determines the corresponding weights of each indicator. Based on field observations and interviews, children's time-related behaviours in learning and daily routines were investigated to identify high-frequency tasks and refine assessment criteria. Consequently, a practical evaluation tool titled "My Schedule" was developed, integrating perception, planning, and regulation tasks to measure children's time management competence. Experimental validation demonstrated that the tool possesses good reliability and validity, and effectively differentiates children's performance across various age and ability levels. The results confirm that this FAHP-based system can objectively assess children's abilities in time perception, task planning, and self-regulation. The study not only enriches the methodological framework for time management assessment but also provides valuable guidance for teachers and parents in cultivating children's executive functions and self-management skills during early education.

Keywords: Time Management Ability, Evaluation System, Assessment Tool, FAHP, Early Childhood Development

1. Introduction

With the release of the Learning and Development Guidelines for Children Aged 3-6, cultivating children's time management skills has gained increasing attention. Effective time management not only impacts learning and daily efficiency but also relates to self-regulation and the development of responsibility. However, surveys indicate that most preschoolers lack a clear sense of time, overly relying on adult reminders, which affects learning outcomes and increases pressure during school transition(Chen, 2024). While existing research focuses on cognitive development and educational practices, developing some assessment tools, these primarily rely on self-report scales. Lacking experimental validation, they struggle to objectively reflect children's actual abilities. A particular research gap exists in time management evaluation tools tailored for children transitioning from preschool to elementary school. To address this, this study introduces the Fuzzy Analytic Hierarchy Process (FAHP). By constructing a three-tiered evaluation model for children's time management abilities, it tackles the challenge of scientifically quantifying and systematically assessing these skills. Based on this model, a task-based evaluation tool was designed. Through experiments, behavioural data on children's task selection, prioritization, and scheduling were collected, quantified, and comprehensively analyzed to validate the tool's effectiveness. The findings establish clear quantitative standards and empirical evidence for children's time management abilities, providing valuable references and support for scientific cultivation in preschool education and the design of related smart products.

2. Literature Review

2.1 Current Status of Children's Time Management

Time management is the core competency through which individuals achieve goals by planning, regulating, and executing activities within limited time frames. It not only impacts learning and work efficiency but also serves as a vital safeguard for self-regulation and mental health(García-Ros & Pérez-González, 2012). For children transitioning from preschool to elementary school, developing foundational time awareness and task planning abilities is a critical prerequisite for adapting to the demands of primary education. Children at this stage are in a critical period of rapid cognitive and behavioural development. Good time management habits help establish regular routines, improve task completion rates, and foster responsibility and independence.

However, research indicates that during the transition from preschool to elementary school, only about 30% of children can independently organize and effectively manage their time. Most lack planning skills and a sense of responsibility, have a vague understanding of time, and exhibit insufficient awareness of self-regulation and planning. They often rely excessively on reminders from parents or teachers and are prone to procrastination(Fu, 2025). Furthermore, differences in educational backgrounds and parenting philosophies across families lead to significant disparities in parental understanding and guidance regarding children's time management: some parents lack scientific time management awareness and struggle to provide timely, effective guidance; others tend to impose adult-oriented, high-intensity schedules, leaving children little room for self-regulation.

These variations not only hinder the development of children's time management skills but also result in inconsistent and subjectively biased parental evaluations of their children's performance. From a cognitive development perspective, executive functions encompass inhibitory control, planning, fluency, working memory, and task switching. Children's time management abilities are closely linked to the development of these executive functions(Do & Eoh, 2022). Executive functions emphasize their critical role in consciously regulating thoughts and actions to achieve long-term goals. During the transition from preschool to elementary school, children's performance in time awareness, task planning, and self-regulation fundamentally reflects the external application of their executive functions in real-life contexts(Best & Miller, 2010). Therefore, assessing time management abilities at this stage should not only describe behavioural outcomes but also be understood through the lens of developmental psychology and cognitive control, considering the developmental characteristics of executive functions.

In summary, establishing a scientific, systematic, and standardized evaluation system for children's time management skills during the transition from preschool to elementary school is particularly essential.

2.2 Time Management Skills Assessment

2.2.1 Research Developments Abroad

Overseas, a relatively mature system has been established in the field of time management competency assessment, primarily comprising two types of tools: scales and tasks. Scale-based tools focus on measuring psychological structures, reflecting individuals' cognitive and behavioural tendencies through self-reporting. While operationally straightforward, they exhibit a high degree of subjectivity. Macan's Time Management Behaviour Model divides time management into four processes: goal setting, scheduling, monitoring, and feedback. Based on this model, the TMB scale was developed to assess individuals' time management behaviours across four dimensions: time perception, task planning, prioritization, and self-regulation(Macan et al., 1990). Britton's Time Management Questionnaire (TMQ), grounded in self-regulation theory, evaluates individuals' time allocation in academic and daily life across three levels: time attitudes, short-term planning, and long-term planning(Britton & Tesser, 1991). The Time Structure Questionnaire (TSQ) examines the level of structuring in time use and goal-directed cognitive processes through six factors: sense of purpose, structured routines, present orientation, effective organization, and persistence(Bond & Feather, 1988). For children, the University of Gothenburg in Sweden employs a time cognition model to evaluate time perception, orientation, and management abilities, particularly suited for children with learning disabilities(Janeslätt et al., 2008).The Time-S Scale focuses on children's planning and time awareness development(Sköld & Janeslätt, 2017) .Despite extensive research on these scales, most studies concentrate on dimensionality and statistical validation, with the theoretical foundations of questionnaire design requiring further refinement.

Task-based tools emphasize behavioural assessment, typically grounded in executive function theory and the Theory of Planned Behaviour. While offering high objectivity, they are operationally complex and costly. The “Plan-a-Day” task requires participants to create daily activity plans within simulated work scenarios, adhering to constraints such as time, location, and duration, thereby evaluating planning abilities at the executive function level (Holt et al., 2011). The BADS-C Zoo Map Task requires children to plan a visit route avoiding redundant paths, with penalties for redundant or ineffective steps, and is widely used in assessing executive function deficits in children(Ballhausen et al., 2017). The “Cooking Breakfast” task requires participants to prepare breakfast while setting the table, assessing planning and execution abilities by recording cooking time differences and multitasking performance (Craik & Bialystok, 2006). Additionally, the “My Schedule” task, grounded in behavioural decision theory, evaluates individuals' task scheduling and time allocation strategies by simulating activity choices in an office environment(Romero et al., 2023). While these tasks provide detailed operational procedures and validation methods, further refinement is possible in areas such as theoretical model selection, key parameter settings, and the correspondence between assessment metrics.

2.2.2 State of China Studies

Domestic research began relatively late and has primarily relied on questionnaire surveys. Huang Xiting and Zhang Zhijie developed the Adolescent Time Management Disposition Scale (ATMD) based on the Time Management Questionnaire (TMQ), Time Management Beliefs (TMB), and Time Self-Efficacy Questionnaire (TSQ), incorporating China's cultural context. This scale assesses adolescents' time management characteristics across three dimensions: time value perception, time monitoring perspective, and time self-efficacy(Huang & Zhang, 2001). Research on time management abilities across age groups in China has largely relied on this questionnaire. However, its applicability to children in the early elementary transition phase lacks scientific validation, and experimental or behavioural assessment methods for this group remain largely unexplored.

Overall, measurement-based tools focus on assessing psychological structures related to children's time management, exhibiting strong structural characteristics. Task-based tools, meanwhile, reflect children's planning and execution abilities in practical operations through contextualized behavioural assessments, demonstrating high ecological validity. However, both types of tools face certain limitations in their theoretical frameworks. Concurrently, research on child development measurement indicates that single assessment methods often struggle to comprehensively capture the developmental characteristics of complex abilities. Particular limitations persist in integrating multidimensional behavioural indicators and addressing developmental stage differences.(Toplak et al., 2013)

Therefore, based on the multidimensional, phased, and contextual assessment principles emphasized in child development measurement, this study draws upon the structured measurement approach of scales and task-based behavioural assessment methods. It employs the Fuzzy Analytic Hierarchy Process (FAHP) to construct a scientific evaluation system for time management abilities in children during the transition from preschool to elementary school. FAHP integrates the structured decision logic of the Analytic Hierarchy Process (AHP) with the membership degree principle of fuzzy mathematics. It hierarchically decomposes dimensions such as time perception, task planning, and self-regulation. Expert judgment addresses ambiguity and uncertainty among indicators, standardizes evaluation weights, and enhances the scientific rigor and objectivity of the assessment system. Building upon this foundation, the study designed the “My Schedule” time management assessment tool for children. This tool presents tasks of varying types and points values, requiring children to independently select and prioritize them. Predefined rules trigger feedback, with task completion scores reflecting their time management proficiency.

3. Method

This study employs a comprehensive approach integrating the Fuzzy Analytic Hierarchy Process (FAHP), expert scoring, experimental validation, and usability assessment, conducted in four steps:

(1) Constructing a three-level evaluation model based on FAHP and calculating the weight of each dimension;

(2) Identifying core tasks and refining scoring criteria by screening high-frequency children's activities through literature analysis, interviews, and expert discussions;

(3) Determined task and rule scores by integrating the model with expert ratings to develop the “My Schedule” tool;

(4) Validated the tool's effectiveness through experimental testing.

3.1 Three-Level Model Construction and Weight Calculation

FAHP is an optimization method developed by Saaty by integrating Fuzzy Comprehensive Evaluation (FCE) with the AHP (Van Laarhoven & Pedrycz, 1983). Children's time management abilities encompass a complex concept involving multiple cognitive dimensions and behavioural elements, such as time perception, task prioritization, and task balancing. They are also influenced by various variables including task difficulty, required time, and completion priority. These factors are difficult to measure precisely with a single numerical value, and assessments by parents or teachers may also carry subjective variations. The organic integration of AHP and FCE leverages the strengths of hierarchical decision analysis and fuzzy comprehensive processing, serving as a crucial technical foundation for the assessment tool of children's time management abilities in this study.

First, AHP is highly suitable for analyzing structured and hierarchical problems. The assessment of children's time management abilities constitutes a comprehensive, multi-level system involving indicators ranging from overarching goals to specific behavioural manifestations. The AHP method breaks down complex evaluation problems into target levels, criterion levels, and indicator levels, establishing a hierarchical structural model. This clearly reveals causal relationships and weight distributions across levels, thereby enhancing the system's scientific rigor and systematicity.

Second, the Fuzzy Comprehensive Evaluation (FCE) method addresses ambiguity and uncertainty inherent in evaluation indicators. Children's time management behaviours in daily life often involve both subjective perceptions and objective manifestations, inevitably introducing fuzzy interference during assessment. The FCE method leverages fuzzy mathematics principles to comprehensively account for the ambiguity inherent in expert subjective evaluations. It normalizes fuzzy judgment matrices and then employs weighted averaging to derive more objective and stable indicator weights. In this study, the introduction of FCE not only helps reduce the bias of subjective scoring but also yields more credible and interpretable comprehensive evaluation results within the complex environment of multidimensional indicator interactions.

3.1.1 Three-Tier Model Construction

(1) Target Laye: children's version of "My Schedule" (A), providing a comprehensive evaluation of children's time management abilities.

(2) Criterion Layer: including task attributes (B) that reflect the intrinsic characteristics of the task, and reasonableness of arrangement (C) that reflects the coordination of task allocation and the distribution of time and effort.

(3) Sub-criterion Layer: task attributes include task importance (b1), which reflects the value of the task in relation to the child's growth or life; task difficulty (b2), which reflects the skills and attention required to complete the task; and time cost (b3), which reflects the time and effort required to complete the task. Task scheduling rationality includes time allocation (c1), which reflects whether tasks are scheduled during appropriate time periods; task sorting (c2), which reflects whether the plan balances work and rest; and task balance (c3), which reflects the coverage of activity types.

3.1.2 Establish a Fuzzy Judgment Matrix

To incorporate fuzziness into the evaluation process, this study employs a fuzzy complementary judgment matrix (0.1–0.9 scaling method), as shown in Table 1, pairwise comparisons are conducted between each element in the criterion layer and sub-criterion layer to quantify their relative importance, forming a fuzzy complementary judgment matrix $R=(a_{ij})n\times n$, namely:

$$R = (a_{ij})n \times n = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad (1)$$

The matrix must satisfy the following conditions(Kubler et al., 2016):

(1) $a_{ij} = 0.5$ (i, j =1, 2, ..., n), the two elements are equally important;

(2) $a_{ij} + a_{ji} = 1$. If $a_{ij} < 0.5$, a_j is more important; otherwise, a_i is more important.

Table 1. 0.1–0.9 scaling method

Scale (a_{ij})	Definition	Explanation
0.5	Equally important	a_i is of equal importance to a_j.
0.6	Slightly important	a_i is slightly more important than a_j.
0.7	Obviously important	a_i is obviously more important than a_j.
0.8	Very important	a_i is much more important than a_j.
0.9	Extremely importan	a_i is far more crucial than a_j.
0.1,0.2,0.3,0.4	Anti-comparison	$a_{ij} = a_i/a_j$, $a_{ji}=a_j/a_i$, $a_{ij}=1-a_{ji}$

3.1.3 Expert Scoring

This study invited 11 experts to form a panel for pairwise comparisons of each factor. The panel comprised 3 elementary school teachers with formal training in childhood education (average teaching experience: 7 years), 4 designers specializing in education or children's product design (average professional experience: 4 years), and 4 parents with extensive experience overseeing children's daily routines. Teachers and designers primarily assessed indicator weightings from pedagogical and functional perspectives, while parents provided supplementary insights based on real-world usage scenarios to enhance the practical applicability of evaluation outcomes.

Each expert assigned scores $a_{ij}^{(k)}$ (k=1, 2,11) using a 0.1–0.9 scale. The geometric mean method was employed to integrate the expert ratings, yielding the final fuzzy complementary judgment matrix aij, calculated as follows:

$$a_{ij} = \sqrt[11]{a_{ij}^{(1)} \times a_{ij}^{(2)} \times \cdots \times a_{ij}^{(11)}} \quad (2)$$

Based on the above method, calculations are performed and uniformly rounded to three decimal places, yielding the fuzzy decision matrix:

the criterion layer is $R_1=\begin{bmatrix} 0.500 & 0.536 \\ 0.464 & 0.500 \end{bmatrix}$;

the task attributes B is $R_2=\begin{bmatrix} 0.500 & 0.603 & 0.589 \\ 0.397 & 0.500 & 0.518 \\ 0.411 & 0.482 & 0.500 \end{bmatrix}$;

the reasonableness of arrangement C is $R_3=\begin{bmatrix} 0.500 & 0.516 & 0.429 \\ 0.484 & 0.500 & 0.449 \\ 0.571 & 0.551 & 0.500 \end{bmatrix}$.

3.1.4 Weight Calculation

Let R = a_{ij} (n×n) denote the fuzzy complementary judgment matrix, and W = (W_1, W_2, W_3, ..., W_n) denote the weight vector. According to the formula in the literature (Huang & Yan, 2025), the weights of each factor can be calculated using the following formula:

$$W_i = \frac{\sum_{i,j=1}^{n} a_{ij} + \frac{n}{2} - 1}{n(n-1)} \qquad (3)$$

The weights for the main criterion layer are W= (0.518, 0.482); the weights for the sub-criterion layers are W_1= (0.365, 0.319, 0.316) and W_2= (0.324, 0.322, 0.354).

3.1.5 Consistency Test

To validate the rationality of the weights, the consistency index I(A,W*) was calculated based on the literature (Zheng et al., 2025). When I (A, W*) ≤ α (typically α = 0.1), the matrix is considered to possess satisfactory consistency. A smaller threshold indicates stricter consistency requirements.

$$I(A,W^*) = \frac{\sum_{i,j=1}^{n} |a_{ij} + b_{ji} - 1|}{n^2} \qquad (4)$$

$$W_{ij} = \frac{w_i}{w_i + w_j} \qquad (5)$$

$$A=(a_{ij})n\times n,\ W^*=(b_{ij})n\times n \qquad (6)$$

The results show that I (R, W*) = 0.009 ≤ 0.1, I (R_1, W_1^*) = 0.003 ≤ 0.1, and I (R_2, W_2^*) = 0.020 ≤ 0.1, all meeting the consistency requirements. After aggregating the weights of the criterion layer and sub-criterion layer, the comprehensive weights for each factor are obtained in Table 2.

Table 2. Summary of Factor Weights

Criterion Layer	Weight	Sub-criterion Layer	Weight
Task Attributes B	0.518	Task Importance b1	0.365
		Task Difficulty b2	0.319
		Time Cost b3	0.316
Reasonableness of Arrangement C	0.482	Time Allocation c1	0.324
		Task Sorting c2	0.322
		Task Balance c3	0.354

Due to space limitations, the complete fuzzy judgment matrices and intermediate calculation steps are provided in Appendix A.

3.2 Form Construction and Task/Rule Scoring Determination

3.2.1 Task and Rule Determination

Early childhood daily activities can be categorized into five types: daily living, learning, play, physical exercise, and others(Zhu, 2024), encompassing routines, assignments, creative thinking, physical movement, and social recreation. Through literature review and interviews, this study identified children's high-frequency daily activities. Following expert deliberation, 11 core tasks were selected, as shown in Figure 1. Selection criteria included activity frequency, habit-forming potential, promotion of cognitive or physical development, suitability for independent completion by children, and ease of parental supervision. These criteria ensured the tasks' representativeness and practicality.

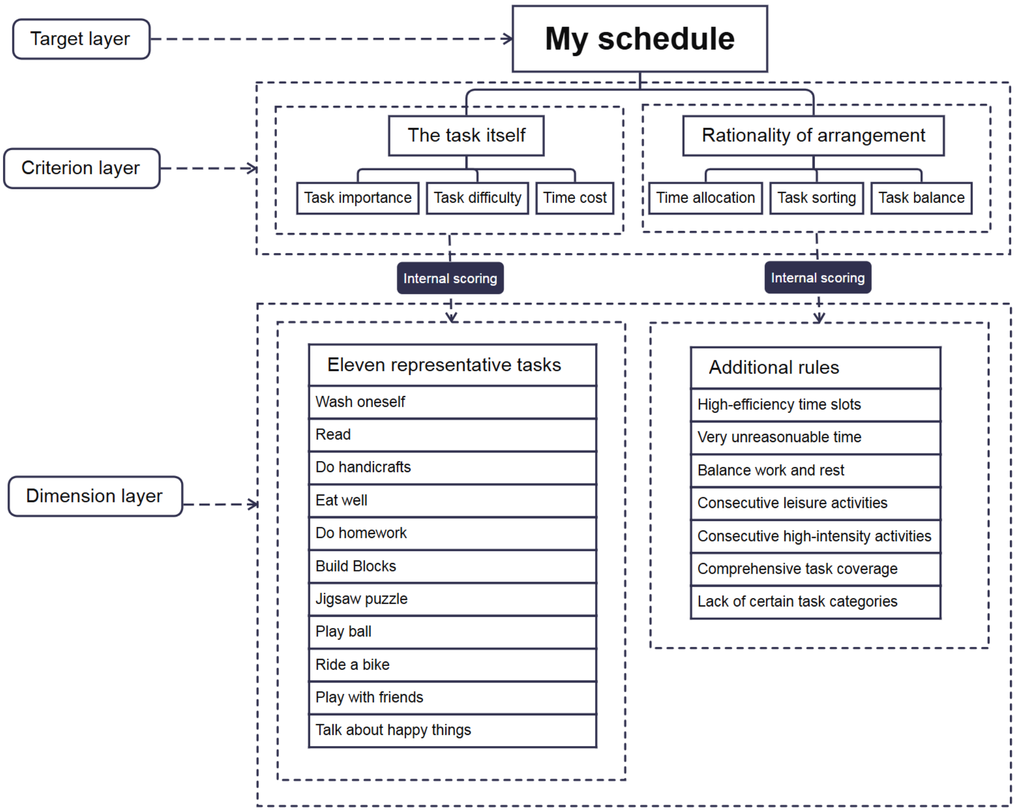

Figure 1. Scoring Method for "My Schedule"

To scientifically evaluate children's time perception, task planning, and self-regulation abilities in practice, this study developed seven supplementary rules under the sub-criteria of "reasonableness of arrangement" within the framework of task sequencing, time allocation, and task balance. As shown in Figure 1. These rules integrate time management and cognitive load theory to standardize and guide the rationality and feasibility of children's learning and life planning.

Regarding time allocation, research indicates that children exhibit peak attention and learning efficiency between 9:00 AM and 4:00 PM(Batejat et al., 1999). Thus, scheduling learning tasks during these hours earns bonus points, while scheduling learning or physical activities late at night incurs penalty points for violating natural rhythms.

Regarding task sequencing, consecutive high-intensity activities deplete cognitive resources, while alternating study with leisure helps sustain focus (Sweller, 2011). Thus, alternating academic tasks with games or sports demonstrates "work-rest balance" and earns points; scheduling multiple consecutive academic or sports tasks, or extended recreational activities, triggers deductions for "high intensity" or "excessive leisure" respectively.

Regarding task balance, children's development requires a mix of learning, daily life, physical activity, and social interaction. When tasks encompass multiple categories, it demonstrates well-rounded development and earns points; if any category is missing, it indicates a one-dimensional structure and triggers point deductions.

3.2.2 Determination of Scores

Based on the established weights for sub-criteria, the aforementioned expert panel was invited to participate in scoring once more. Each specific task and supplementary rule were calibrated using a 1–5 point scale under its corresponding sub-criteria dimension. This scoring process aimed to characterize the attribute levels of different tasks and rules across sub-criteria dimensions, rather than re-evaluating the importance of the sub-criteria themselves.

To mitigate the impact of subjective judgment differences among individual experts, the arithmetic mean method was employed to consolidate the scoring results from multiple experts, yielding internal scores for each task or rule within their corresponding sub-criteria dimensions. Subsequently, these internal scores were weighted using the sub-criteria layer weights established in the three-level analysis model to generate comprehensive scores.

Let the overall weight of a sub-criterion layer be N, and the geometric mean of the internal scores for the corresponding tasks or rules be K. Then the overall score S for that task or rule is expressed as:

$$S_{task} = N_{b1}\times K_{b1}+N_{b2}\times K_{b2}+N_{b3}\times K_{b3} \tag{7}$$

$$S_{rule} = N_{c1}\times K_{c1}+N_{c2}\times K_{c2}+N_{c3}\times K_{c3} \tag{8}$$

Among these, b1、b2 and b3 represent the three sub-criteria layer of task importance、task difficulty and time cost, respectively. c1、c2 and c3 represent the three sub-criteria layer of time allocation、task sorting and task balance, respectively. Ultimately, the comprehensive scoring results for each task and additional rule are shown in Tables 3 and 4.

Table 3. Task Attributes

Task Attributes	**Task**	**Sub-Criterion Level**	**Internal scoring**	**Overall Score**
Daily living	Wash oneself	Task Importance	3.727	1.423
		Task Difficulty	2.455	
		Time Cost	1.909	
	Eat well	Task Importance	4.545	1.861
		Task Difficulty	3.000	
		Time Cost	3.091	
Learning	Read	Task Importance	4.727	2.150
		Task Difficulty	3.818	
		Time Cost	3.818	
	Do handicrafts	Task Importance	3.455	1.879
		Task Difficulty	3.727	
		Time Cost	3.727	
	Do homework	Task Importance	3.909	2.055
		Task Difficulty	4.182	
		Time Cost	3.818	
Play	Build Blocks	Task Importance	3.091	1.571
		Task Difficulty	3.000	
		Time Cost	3.000	

Task Attributes	Task	Sub-Criterion Level	Internal scoring	Overall Score
	Jigsaw puzzle	Task Importance	2.909	1.761
		Task Difficulty	3.636	
		Time Cost	3.727	
Physical exercise	Play ball	Task Importance	3.818	1.723
		Task Difficulty	2.909	
		Time Cost	3.182	
	Ride a bike	Task Importance	3.545	1.791
		Task Difficulty	3.182	
		Time Cost	3.636	
Others	Play with friends	Task Importance	4.091	1.595
		Task Difficulty	2.000	
		Time Cost	3.000	
	Talk about happy things	Task Importance	4.455	1.709
		Task Difficulty	2.636	
		Time Cost	2.636	

Table 4. Reasonableness of Arrangement

Reasonableness of Arrangement	Rule	Sub-Criterion Level	Internal scoring	Overall Score
Time Allocation	High-efficiency time slots (Bonus points)	Time Allocation	4.091	1.676
		Task Sorting	3.091	
		Task Balance	3.273	
	Late-night time slots (Deduction item)	Time Allocation	3.364	1.324
		Task Sorting	2.455	
		Task Balance	2.455	
Task Sequencing	Balance work and rest (Deduction item)	Time Allocation	3.545	1.783
		Task Sorting	3.818	
		Task Balance	3.727	
	Consecutive leisure activities (Deduction item)	Time Allocation	2.636	1.197
		Task Sorting	2.364	
		Task Balance	2.455	
	Consecutive high-intensity activities (Deduction item)	Time Allocation	2.636	1.343
		Task Sorting	3.000	
		Task Balance	2.727	
Task Balance		Time Allocation	3.727	1.843
		Task Sorting	3.727	

Reasonableness of Arrangement	Rule	Sub-Criterion Level	Internal scoring	Overall Score
	Comprehensive task coverage (Bonus points)	Task Balance	4.000	
	Lack of certain task categories (Deduction item)	Time Allocation	2.364	1.262
		Task Sorting	2.455	
		Task Balance	3.000	

To ensure comparability between different tasks and rules, the composite scores undergo linear normalization:

$$S_n^{norm} = \frac{S_n - S_{min}}{S_{max} - S_{min}} \quad (9)$$

Further map the normalized scores to integer values to form scoring rules directly usable by children:

$$P_n = [S_n^{norm} \times (P_{max} - P_{min})] \quad (10)$$

Where Pn denotes the final integer score, Pmin and Pmax represent the minimum and maximum scores, respectively.

Based on the aforementioned expert weighting and internal scoring results, the scoring rules directly applicable to children can be derived, as shown in Table 5.

Table 5. Scoring Criteria

Category	Task	Points
Daily living	Wash oneself	3
	Eat well	4
Learning	Read	5
	Do handicrafts	4
	Do homework	5
Play	Build Blocks	3
	Jigsaw puzzle	4
Physical exercise	Play ball	4
	Ride a bike	4
Others	Play with friends	3
	Talk about happy things	4
Additional Rule: The same rule cannot be applied cumulatively; it may be applied only once.		
Bonus Points: High-efficiency time slots (academic tasks around 9:00/16:00) +3 Balance work and rest +4 Comprehensive task coverage +4		
Penalty Points: Very unreasonable time (disrupting circadian rhythm) -2 Consecutive leisure activities (entertainment ≥3) -2 Consecutive high-intensity activities (learning /physical exercise ≥3) -2 Lack of certain task categories -2		

3.3 Evaluation Tool for Time Management Skills in Children During the Transition from Preschool to Elementary School

The experimental evaluation tool "My Schedule" designed in this study simulates children's daily task arrangements to assess their time management abilities. As shown in Figure 2, the tool comprises task cards (11 cards covering activities such as daily life, study, play, and sports), a schedule divided into morning/afternoon/evening segments, and a scoring rules table, as shown in Table 5. During the experiment, children independently selected 8 tasks from the task cards and allocated them to specific time slots according to the rules. Researchers calculated scores based on task arrangements and rules to quantify children's time perception, task planning, and self-regulation abilities, providing data for subsequent analysis.

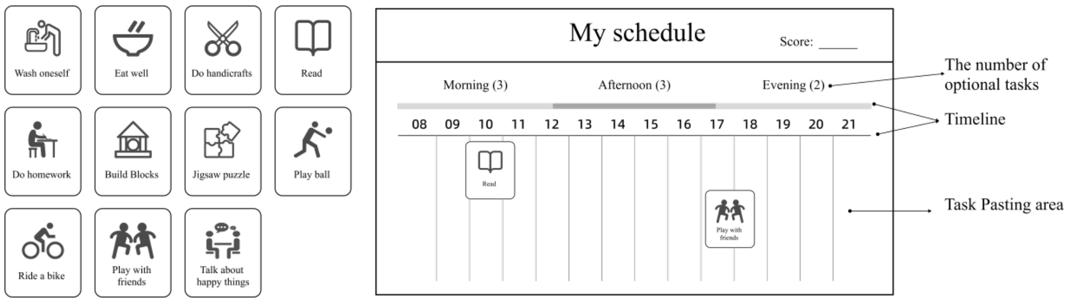

Figure 2. Instructions for Using the 11 Task Cards and Schedule

3.3.1 Classification of Children's Time Management Competency Levels

Children can independently select 8 tasks from "My Schedule", with each additional rule triggering a maximum of one time. Task scores range from 29 to 34 points, while rule scores range from 8 to 11 points. The total score is calculated as the sum of task and rule scores, ranging from 21 to 45 points. To facilitate the quantification of children's time management abilities and enable comparisons among different children, the total score can be converted into a capability percentage:

$$\text{Ability Percentage} = \frac{S_{\text{total score}} - 21}{45 - 21} \times 100\% \tag{11}$$

Based on ability percentages, children's time management skills are categorized into five levels, as shown in Table 6.

Table 6. Children's Time Management Ability Rating Scale

Proficiency Level	Percentage	Score Range	Explanation
High Level	≥80%	41～45	Task selection is reasonable and scheduling is balanced
Mid-High Level	65%-79%	37～40	Most tasks are reasonable, and scheduling is generally appropriate
Mid-Level	50%-64%	33～36	Task selection requires optimization, and scheduling is uneven
Mid-Low Level	35%-49%	30～32	Task selection is unreasonable, and planning is inadequate
Low Level	＜35%	21～29	Task selection and scheduling are clearly unreasonable

4. Experimental Verification

To validate the effectiveness of "My Schedule", this study conducted a small-scale preliminary experiment to assess its validity and usability in evaluating children's time management skills.

4.1 Participants and Recruitment

The experiment recruited 30 healthy children aged 5 – 9 years in Hangzhou, China (M = 7.0), with an equal number of boys and girls. None of the children had known cognitive or learning disabilities. Parents or teachers served as external raters. This age group represents the transition from preschool to elementary school and possesses basic self-expression and form-filling abilities, making it suitable for the preliminary validation of the "My Schedule" tool. A sample size of 30 was adopted to provide initial evidence of feasibility, validity, and usability.

Prior to participation, parents or legal guardians were provided with detailed information about the study's objectives and procedures and gave written informed consent. Participation was voluntary, and no missing data were observed. Potential confounding factors, such as prior exposure to time-management activities and parental involvement, were recorded but not statistically controlled. Age-related effects were not examined in this study and will be explored in future research with larger samples.

The study involved minimal risk and focused on evaluating an educational tool. In line with institutional guidelines, formal ethical approval was considered exempt. All procedures were conducted in accordance with relevant ethical standards for research involving human participants.

4.2 Data Collection

This study not only records children's actual operational scores S to reflect their time management competency levels but also collects parent/teacher percentage ratings C and overall performance scores P to validate the instrument's criterion-related validity and provide comparative references. The percentage rating C quantifies children's performance across the four core dimensions of the TMB scale, offering multidimensional behavioural references for S. The overall performance score P evaluates children's comprehensive abilities in both the operational process and outcomes from a macro perspective, enhancing the scientific rigor and reliability of comparisons between S and parental ratings. Concurrently, children completed the System Usability Scale (SUS)(Bangor et al., 2008) to assess the tool's usability, efficiency, and subjective satisfaction, ensuring the operational feasibility and applicability of "My Schedule" within the target population.

(1) Children's "My Schedule" Score S. Children independently complete the schedule based on daily habits and comprehension, with researchers calculating the score accordingly.

(2) Parent Percentage Score C. Percentage scores are assigned based on the four dimensions of the TMB scale. The average of these four dimensions is converted into a simulated score P on the same scale as S, calculated as $P = S_{total\ scoe\ max} \times$ (Dimension 1 + Dimension 2 + Dimension 3 + Dimension 4) / 4, serving as the criterion reference.

(3) Parent Performance Score P. Scores children's overall performance in completing the schedule (maximum 45 points) to evaluate comprehensive ability in both process and outcome.

(4) SUS Score. The SUS scale was simplified into a 5－point rating scale suitable for children aged 5–9. Researchers provided guidance to ensure each child understood the meaning of each item and could respond accurately. Children completed the scale based on their actual usage experience to assess the tool's usability and learnability.

4.3 Experimental Procedure

Children completed the “My Schedule” and SUS scales under guidance, while parents filled out the rating forms. The experiment included task instruction, hands-on practice, scoring, and interviews, while observing children's time management behaviours and collecting evaluation data, as shown in Figures 3 and 4.

Figure 3. The Process of Children Filling Out “My Schedule”

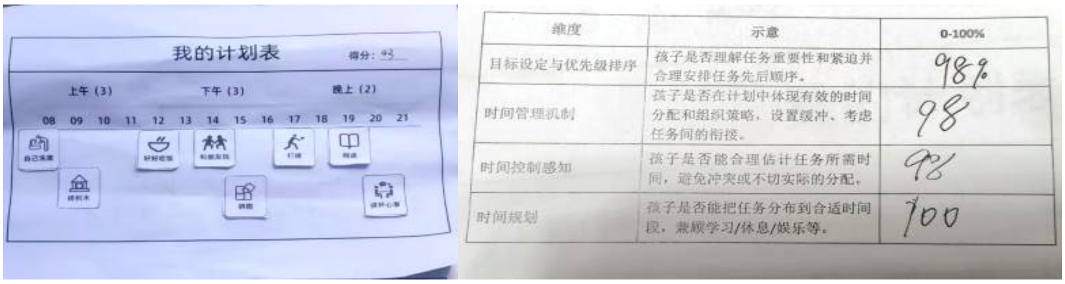

Figure 4. Children's Completed “My Schedule” and Parent Scores

4.4 Results and Discussion

4.4.1 Criterion-Related Validity Verification

Experimental data indicate that the overall trends for children's “My Schedule” scores (S), parental percentage ratings (C), and performance ratings (P) are consistent, as shown in Figure 5, suggesting that this tool possesses a certain degree of criterion-related validity in reflecting children's time management abilities. Overall, children's abilities span all proficiency levels, though predominantly at the medium to upper-medium range. However, significant discrepancies exist between S, C, and P scores for certain children (e.g., Child 11, Child 19). This variation likely stems from differing evaluation

perspectives and scoring dimensions: S emphasizes children's practical operational performance within the system, C focuses on comprehensive assessment across dimensions, while P evaluates overall task completion effectiveness. Further analysis indicates that this primarily stems from some parents' tendency to overestimate their children's time management skills, reflecting insufficient recognition of children's actual capabilities in this area and resulting in individual biases.

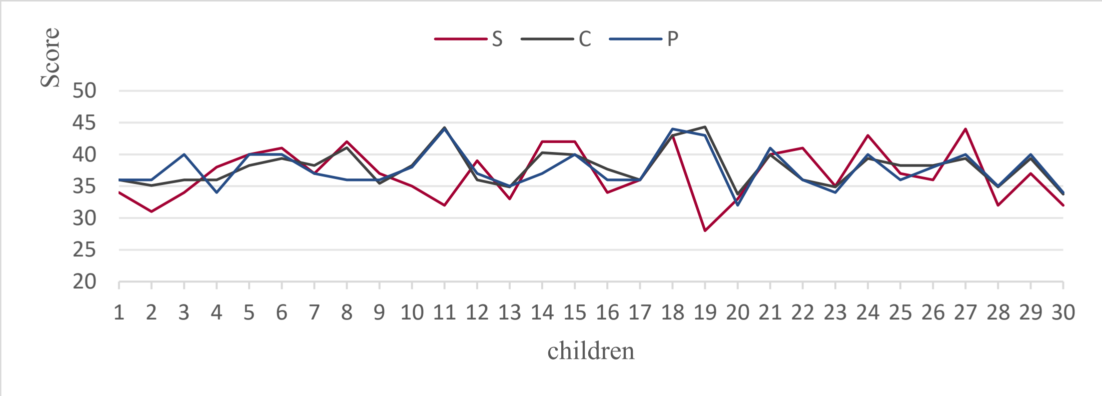

Figure 5. S、 C、P Scores

Table 7. Paired T-test Results

Comparison	Mean Difference	SD	t	df	p	Cohen's d	95% CI
S VS C	−0.990	4.367	−1.247	29	0.222	−0.227	−2.625 ~ 0.636
S VS P	−0.770	4.554	−0.922	29	0.364	−0.169	−2.467 ~ 0.934
C VS P	0.230	1.651	0.755	29	0.457	0.139	−0.389 ~ 0.844

To further examine the consistency among the three measures, paired-samples t-tests were conducted, as shown in Table 7. Results showed that the mean "My Schedule" score was S = 39.933 (SD = 4.201), compared with parental percentage ratings C = 37.928 (SD = 2.863) and performance ratings P = 37.692 (SD = 3.052). No statistically significant differences were found between S and C ($t(29) = -1.247$, $p = 0.222$) or between C and P ($t(29) = 0.755$, $p = 0.457$).

Effect size analysis further supported these findings. The paired-sample Cohen's d values were -0.227 for S vs. C、−0.169 for S vs. P and 0.139 for C vs. P, indicating small to negligible practical differences between children's system-based scores and parent-based evaluations. Taken together, these results suggest a high degree of consistency among the three measures and provide empirical evidence supporting the criterion-related validity of the "My Schedule" assessment tool.

4.4.2 Availability Verification

In the SUS scoring, 68 points is generally considered the average usability benchmark, with scores above 80 indicating higher usability, while scores below 50 suggest significant issues. Specifically, items 4 and 8 assess learnability, while the remaining items evaluate overall usability. As shown in Table 8, statistical results indicate that the mean SUS total score for this study's sample was 81.083 (>80), with mean scores of 79.583 for learnability and 81.458 for usability. Both scores fall within the higher range, demonstrating that this tool offers a positive user experience and effective usability.

To further quantify the statistical significance of system usability, an effect size and confidence interval analysis was conducted on the SUS total score. Results indicate (Table 9) that the 95% confidence interval for the mean total score is [78.096, 84.071], with Cohen's d = 1.57, reaching a large effect size. This demonstrates that the system not only exhibits a high mean usability score but also possesses statistically significant and robust stability in its effects. To validate the robustness of the results, this study conducted a sensitivity analysis. The findings indicate that the overall SUS score remains unaffected by outliers or subgroup divisions, demonstrating the stable and reliable usability of the system.

Table 8. System Usability Scale Scores

Q1	4.067
Q2	1.667
Q3	4.267
Q4	1.900
Q5	4.433
Q6	1.600
Q7	3.867
Q8	1.733
Q9	4.367
Q10	1.667
SUS	81.083
Learnability	79.583
Usability	81.458

Table 9. SUS scale analysis and confidence interval calculation

	Mean ± SD	**Mean 95% CI(LL)□**	**Mean 95% CI(UL)□**
Q1	4.067±0.691	3.819	4.314
Q2	1.667±0.802	1.380	1.954
Q3	4.267±0.640	4.038	4.496
Q4	1.900±0.885	1.583	2.217
Q5	4.433±0.679	4.190	4.676
Q6	1.600±0.563	1.398	1.802
Q7	3.867±0.819	3.573	4.160
Q8	1.733±0.691	1.486	1.981
Q9	4.367±0.765	4.093	4.640
Q10	1.667±0.606	1.450	1.884
Total Score	81.083±8.348	78.096	84.071

4.4.3 Interview Findings

To gain deeper insights into the tool's applicability, semi-structured interviews were conducted with children and parents, focusing on consistency, practicality, and improvement suggestions.

Consistency: Most parents reported that children's self-assessments on the task sheet generally aligned with their daily behavioural performance, with assigned tasks primarily involving routine activities. However, instances occurred where parental ratings at the individual level did not fully match children's scores, indicating that experimental scores and peer evaluations may be influenced by observational perspectives and contextual factors.

Practicality: Parents generally agreed the tool helped children clearly articulate and organize daily tasks, facilitating assessment of their time planning and execution abilities. Children reported ease of use, understanding task card meanings, and completing basic arrangements, demonstrating the tool's strong applicability in promoting proactive task planning and execution.

Improvement Suggestions: Some children noted task cards were too large and did not fully align with time scales. Parents pointed out the timeline was insufficiently prominent, potentially hindering children's intuitive understanding of task scheduling. Recommendations include optimizing card dimensions and enhancing the visual clarity of the timeline to further improve the tool's operational convenience and adaptability.

Overall, "My Schedule" demonstrated good effectiveness and usability in preliminary testing. It effectively reflects children's time management levels to a certain extent and has gained widespread recognition from both parents and children. Simultaneously, the improvement suggestions provided by parents offer valuable reference directions for further refining the tool.

5. Conclusion

This study explored assessment methods for time management abilities in children aged 5 to 9, using "My Schedule" as the core tool. Findings indicate that children can comprehend task cards and complete daily task arrangements. Parent ratings generally align with children's actual performance scores, demonstrating that this tool effectively reflects children's performance in time planning and execution. It exhibits sound criterion validity and practical usability. "My Schedule" not only realistically portrays children's time management proficiency but also provides a practical tool foundation for cultivating and intervening in time management skills during the preschool-to-elementary transition phase. Based on these findings, it is recommended to guide children in conducting self-directed planning and execution activities across multidimensional contexts such as daily life, learning, play, and physical activities. Due to limitations in sample size and methodology, this study was unable to comprehensively examine other forms of validity, including construct validity, known-related validity, and convergent validity with standardized childhood time processing instruments (e.g., Time-S). Future research may enhance the scientific assessment system for children's time management abilities by expanding sample sizes, optimizing task design, incorporating cross-validation through alternative evaluation methods, and progressively establishing normative data.

References

Ballhausen, N., Mahy, C. E., Hering, A., Voigt, B., Schnitzspahn, K. M., Lagner, P., ... & Kliegel, M. (2017). Children's planning performance in the Zoo Map task (BADS-C): Is it driven by general cognitive ability, executive functioning, or prospection?. Applied Neuropsychology: Child. [Online]. 6(2), pp. 138-144. Available: https://doi.org/10.1080/21622965.2015.1124276

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. Intl. Journal of Human-Computer Interaction. [Online]. 24(6), pp. 574-594. Available: https://doi.org/10.1080/10447310802205776

Batejat, D., Lagarde, D., Navelet, Y., & Binder, M. (1999). Evaluation of the attention span of 10,000 school children 8-11 years of age. Archives de pediatrie: organe officiel de la Societe francaise de pediatrie. [Online]. 6(4), pp. 406-415. Available: https://doi.org/10.1016/S0929-693X(99)80222-X

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. Child development. [Online]. 81(6), pp. 1641-1660. Available: https://doi.org/10.1111/j.1467-8624.2010.01499.x

Bond, M. J., & Feather, N. T. (1988). Some correlates of structure and purpose in the use of time. Journal of personality and social psychology. [Online]. 55(2), pp. 321-329. Available: https://doi.org/10.1037/0022-3514.55.2.321

Britton, B. K., & Tesser, A. (1991). Effects of time-management practices on college grades. Journal of educational psychology. [Online]. 83(3), pp. 405-410. Available: https://doi.org/10.1037/0022-0663.83.3.405

Craik, F. I., & Bialystok, E. (2006). Planning and task management in older adults: Cooking breakfast. Memory & Cognition. [Online]. 34(6), pp. 1236-1249. Available: https://doi.org/10.3758/BF03193268

Do, Y. M., & Kyung, E. Y. (2022). The Executive Function of Childhood and Adolescence: A Review for Comprehensive Understanding. *Journal of the Korean Society of Psychotherapy (kap)*, 14(1), pp. 099-122.

Fu, C. (2025). Time management status and ability enhancement strategies of senior kindergarten children from the perspective of preschool–primary school transition. *Exam Weekly*, (2), pp. 156–159.

García-Ros, R., & Pérez-González, F. (2012). Spanish version of the time management behavior questionnaire for university students. The Spanish Journal of Psychology. [Online]. 15(3), pp. 1485-1494. Available: https://doi.org/10.5209/rev_SJOP.2012.v15.n3.39432

Holt, D. V., Rodewald, K., Rentrop, M., Funke, J., Weisbrod, M., & Kaiser, S. (2011). The Plan-a-Day approach to measuring planning ability in patients with schizophrenia. Journal of the International Neuropsychological Society. [Online]. 17(2), pp. 327-335. Available: https://doi.org/10.1017/S1355617710001712

Huang, R., & Yan, B. (2025). Design research on a badminton shuttlecock steamer based on the fuzzy analytic hierarchy process.Design. [Online]. 38(13), pp. 118–122. Available: https://doi.org/10.20055/j.cnki.1003-0069.002789

Huang, X., & Zhang, Z. (2001). Development of the Adolescents' Time Management Disposition Inventory. *Acta Psychologica Sinica*, 33(4), pp. 338–343.

Janeslätt, G., Granlund, M., Alderman, I., & Kottorp, A. (2008). Development of a new assessment of time processing ability in children, using Rasch analysis. Child: care, health and development. [Online]. 34(6), pp. 771-780. Available: https://doi.org/10.1111/j.1365-2214.2008.00865.x

Kubler, S., Robert, J., Derigent, W., Voisin, A., & Le Traon, Y. (2016). A state-of the-art survey & testbed of fuzzy AHP (FAHP) applications. Expert systems with applications. [Online]. 65, pp. 398-422. Available: https://doi.org/10.1016/j.eswa.2016.08.064

Macan, T. H., Shahani, C., Dipboye, R. L., & Phillips, A. P. (1990). College students' time management: Correlations with academic performance and stress. Journal of educational psychology. [Online]. 82(4), pp. 760-768. Available: https://doi.org/10.1037/0022-0663.82.4.760

Romero, M., Hernández, J. M., Juola, J. F., Casadevante, C., & Santacreu, J. (2023). A complementary methodology to assess time management behaviors. Current Psychology. [Online]. 42(3), pp. 1966-1978. Available: https://doi.org/10.1007/s12144-021-01578-x

Sköld, A., & Janeslätt, G. K. (2017). Self-rating of daily time management in children: psychometric properties of the Time-S. Scandinavian journal of occupational therapy. [Online]. 24(3), pp. 178-186. Available: https://doi.org/10.1080/11038128.2016.1185465

Sweller, J. (2011). Cognitive load theory. In Psychology of learning and motivation (Vol. 55, pp. 37-76). Academic Press. [Online]. Available: https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct?. Journal of child psychology and psychiatry. [Online]. 54(2), pp. 131-143. Available: https://doi.org/10.1111/jcpp.12001

Van Laarhoven, P. J., & Pedrycz, W. (1983). A fuzzy extension of Saaty's priority theory. Fuzzy sets and Systems. [Online]. 11(1-3), pp. 229-241. Available: https://doi.org/10.1016/S0165-0114(83)80082-7

Zheng, A., Hu, Z., Xiao, G., Zhang, L., & Liu, Y. (2025). Impact assessment of plateau hydropower development on the survival of endangered reptiles based on the FAHP method. *Hydropower*, pp. 1–7.

Zhu, X. (2024). Cultivating time management ability of senior kindergarten children from the perspective of preschool–primary school transition: A case study of time-list practice activities. Educational Observation. [Online]. 13(36), pp. 108–111. Available: https://doi.org/10.16070/j.cnki.cn45-1388/g4s.2024.36.024

Appendix A: Detailed FAHP Procedures and Calculations

This appendix provides detailed calculations for the Fuzzy Analytic Hierarchy Process (FAHP) employed in the “My Schedule” tool within this study, including: construction of the fuzzy complementary judgment matrix, integration of expert ratings, weight calculation, and consistency testing. The purpose is to ensure methodological transparency and reproducibility.

A1. Full Fuzzy Complementary Judgment Matrices

To quantify the relative importance of evaluation indicators while accounting for the inherent uncertainty in expert judgments, this study employs a fuzzy complementary judgment matrix based on a 0.1–0.9 scale. Pairwise comparisons are conducted between indicators at the criterion and subcriterion levels, and the aggregated expert judgments are used to construct the final matrix.

A1.1 Criterion Layer Matrix (R1)

The criterion layer compares the relative importance of task attributes (B) and reasonableness of arrangement (C) in assessing children's time management abilities.An expert provided the following data example:

	Task Attributes (B)	Reasonableness of Arrangement (C)
B	0.5	0.6
C	0.4	0.5

Diagonal elements equal 0.5, indicating equal importance, while off-diagonal elements satisfy the complementarity condition $a_{ij} + a_{ji} = 1$.

A1.2 Sub-criterion Layer Matrix for Task Attributes (R2)

This matrix represents pairwise comparisons among the three task-related subcriteria: task importance, task difficulty, and time cost. It reflects experts' judgments regarding the relative contribution of each task attribute to children's time management performance.An expert provided the following data example:

	Task Importance (b1)	Task Difficulty (b2)	Time Cost (b3)
b1	0.5	0.7	0.7
b2	0.3	0.5	0.5
b3	0.3	0.5	0.5

A1.3 Sub-criterion Layer Matrix for Reasonableness of Arrangement (R3)

This matrix reflects experts' assessment of scheduling rationality, encompassing time allocation, task sequencing, and task balance. It reflects experts' judgments on the importance of each dimension in evaluating the rationality of task arrangements.An expert provided the following data example:

	Time Allocation (c1)	Task Sequencing (c2)	Task Balance (c3)
c1	0.5	0.6	0.5
c2	0.4	0.5	0.3
c3	0.5	0.7	0.5

A2. Example of Expert Rating Aggregation

To illustrate aggregation, one representative indicator pair (b1 vs b2) is presented. Individual ratings are not fully listed for brevity.

Expert	$a_{ij}^{(k)}$
1	0.6
2	0.7
3	0.3
4	0.6
5	0.7
6	0.7
7	0.7
8	0.6
9	0.7
10	0.5
11	0.7

Aggregated value using geometric mean:

$$a_{ij} = \sqrt[11]{a_{ij}^{(1)} \times a_{ij}^{(2)} \times \cdots \times a_{ij}^{(11)}} = 0.603$$

This method mitigates the impact of extreme judgments and enhances the stability of scoring results.

Following the above method, one can obtain the fuzzy complementary judgment matrix for all indicators.

The fuzzy complementary judgment matrix at the criterion layer is:

	Task Attributes (B)	Reasonableness of Arrangement (C)
B	0.500	0.536
C	0.464	0.500

The fuzzy complementary judgment matrix for the sub-criterion level "task attributes" is as follows:

	Task Importance (b1)	Task Difficulty (b2)	Time Cost (b3)
b1	0.500	0.603	0.589
b2	0.397	0.500	0.518
b3	0.411	0.482	0.500

The fuzzy complementary judgment matrix for the sub-criterion level "reasonableness of arrangement" is as follows:

	Time Allocation (c1)	Task Sequencing (c2)	Task Balance (c3)
c1	0.500	0.516	0.429
c2	0.484	0.500	0.449
c3	0.571	0.551	0.500

A3. Weight Calculation Procedure

Weights are calculated from fuzzy complementary judgment matrices using:

$$W_i = \frac{\sum_{i,j=1}^{n} a_{ij} + \frac{n}{2} - 1}{n(n-1)}, i = 1,2,3, i = 1,2,\dots,n$$

The weights for the main criteria layer and sub-criteria layer are:

Criterion Layer	**Weight**	**Sub-criterion Layer**	**Weight**
Task Attributes B	0.518	Task Importance b1	0.365
		Task Difficulty b2	0.319
		Time Cost b3	0.316
Reasonableness of Arrangement C	0.482	Time Allocation c1	0.324
		Task Sequencing c2	0.322
		Task Balance c3	0.354

A4. Consistency Test Computation

To ensure logical coherence, each matrix was tested for consistency.

(1) Construct the derived matrix:

$$W_{ij} = \frac{w_i}{w_i + w_j}$$

$$A=(a_{ij})n\times n,\ W^{*}=(b_{ij})n\times n$$

(2) Construct the derived matrix:

$$I(A,W^{*}) = \frac{\sum_{i,j=1}^{n} |a_{ij}+b_{ji}-1|}{n^2}$$

Consistency Results is as follows:

	Criterion Layer	**Sub-criterion Layer Task Attributes B**	**Sub-criterion Layer Reasonableness of Arrangement C**
I(A,W*)	0.009 ⩽ 0.1	0.003 ⩽ 0.1	0.02 ⩽ 0.1

All matrices meet the consistency requirement, confirming that derived weights are reliable.