FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

CAPE-KIDS: Cognitive and Physical Evaluation for Kids' Integrated Development and Success

Matheus Rodrigues Felizardo



Mestrado em Engenharia de Software

Supervisor: Prof. António Coelho, FEUP

Supervisor: Nuno Rodrigues, INESC TEC

July 30, 2025

© Matheus Rodrigues Felizardo, 2025

CAPE-KIDS: Cognitive and Physical Evaluation for Kids' Integrated Development and Success

Matheus Rodrigues Felizardo

Mestrado em Engenharia de Software

July 30, 2025

Acknowledgments

I would like to express my deepest gratitude to the people who supported and inspired me throughout this journey.

To my mother, Dora, and my sister, Jessica, for their unconditional love, encouragement, and belief in me every step of the way. To my partner, Josceline, for her support, patience, and understanding during long days and nights of writing and development.

To my supervisors, Nuno Rodrigues and Àntonio Coelho, also to Eva de Oliveira and to Sónia Machado, whose guidance, feedback, and motivation were fundamental throughout the research process. Their contributions were invaluable not only in the academic direction of this dissertation but also in the development and evaluation of the CAPE-KIDS platform and in the preparation of articles for scientific conferences.

I am also grateful to my friends and fellow students, Rangel Soares and Layenis Sobrinho, for their constant encouragement, companionship, and motivation throughout the entire academic journey and especially during the dissertation phase.

This work would not have been possible without the support, care, and inspiration I received from each of you. Thank you.

Matheus Rodrigues Felizardo

Resumo

As funções executivas (FE) — como controlo inibitório, memória de trabalho e flexibilidade cognitiva — são fundamentais para o desenvolvimento infantil e estão associadas a competências escolares, autorregulação e saúde cognitiva. No entanto, persistem desafios significativos quanto à padronização, validade ecológica e acessibilidade das ferramentas utilizadas para a sua avaliação em crianças.
Esta dissertação apresenta o **CAPE-KIDS**, uma plataforma digital orientada para a investigação, desenvolvida com o objetivo de apoiar a criação, execução e análise de tarefas cognitivas e protocolos de treino focados na avaliação de FE em crianças dos 5 aos 9 anos. O projeto foi sustentado por uma revisão de escopo da literatura científica, conduzida de acordo com as diretrizes PRISMA-ScR, que identificou as tarefas mais utilizadas na investigação com esta população.
A plataforma oferece uma arquitetura modular e uma interface visual baseada em timelines, permitindo a criação de estudos personalizados de forma acessível, mesmo para utilizadores sem experiência em programação. A implementação técnica foi realizada com NestJS no *backend* e React no *frontend*. O desenvolvimento contou com a participação de especialistas das áreas da psicologia e neurodesenvolvimento infantil, que contribuíram na análise das tarefas, no desenho da plataforma e na sua validação. Como prova de conceito, foram integradas tarefas amplamente utilizadas na literatura, como a *Go/No-Go* e a *Flanker task*, com suporte para configurações experimentais e recolha de dados sobre o desempenho dos participantes.

Palavras-chave: Funções Executivas, Avaliação Cognitiva, Desenvolvimento Infantil, Engenharia de Software, Plataforma Digital, Interação Humano-Computador, Design de Tarefas, Padronização.

Abstract

Executive functions (EF) — such as inhibitory control, working memory, and cognitive flexibility — are fundamental for child development and are closely associated with school readiness, self-regulation, and cognitive health. However, challenges persist regarding the standardization, ecological validity, and accessibility of tools used to assess EF in children.
This dissertation presents **CAPE-KIDS**, a research-oriented digital platform designed to support the creation, execution, and analysis of cognitive tasks and training protocols focused on EF in children aged 5 to 9 years. The project was guided by a scoping review of the scientific literature, following PRISMA-ScR guidelines, which identified the most commonly used tasks for this age group.
The platform provides a modular architecture and a visual, timeline-based interface, enabling the design of customized studies without requiring programming skills. It was implemented using NestJS for the backend and React for the frontend. The development process involved experts in psychology and child neurodevelopment, who contributed to the analysis of tasks, platform design, and validation. As a proof of concept, the platform includes widely used EF tasks such as *Go/No-Go* and *Flanker task*, with support for experimental configurations and accurate performance data collection.

Keywords: Executive Functions, Cognitive Assessment, Child Development, Software Engineering, Digital Platform, Human-Computer Interaction, Task Design, Standardization.

"And once the storm is over,
you won't remember how you made it through...
But you'll know that you're not the same person who walked in."

Haruki Murakami
(adapted from Kafka on the Shore)

Contents

1 Introduction **1**
- 1.1 Context and Relevance . . . 1
- 1.2 Problem Statement . . . 1
- 1.3 Proposed Solution: The CAPE-KIDS Platform . . . 2
- 1.4 Research Question and Objectives . . . 2
- 1.5 Dissertation Structure . . . 2

2 State of the Art **4**
- 2.1 Executive Functions: Foundational Concepts . . . 4
- 2.2 The State of Digital EF Assessment . . . 5
- 2.3 Review Methodology . . . 5
- 2.4 Key Findings . . . 6
  - 2.4.1 Lack of Standardization Across Studies . . . 7
  - 2.4.2 The Need for a Unified Framework . . . 7
  - 2.4.3 Main Executive Function Tasks Identified . . . 8
- 2.5 Challenges and Opportunities . . . 8
- 2.6 Limitations . . . 9
- 2.7 Conclusion and Implications for Platform Design . . . 10
- 2.8 Chapter Summary . . . 11

3 Research Methodology **12**
- 3.1 Overall Research Approach . . . 12
- 3.2 Platform Development Methodology . . . 13
- 3.3 Validation Methodology . . . 13
  - 3.3.1 Formative Evaluation with Experts . . . 13
  - 3.3.2 Methodology for the Experimental Study with Children . . . 13
- 3.4 Chapter Summary . . . 14

4 System Specification **15**
- 4.1 Requirements . . . 15
  - 4.1.1 User and Participant Management . . . 15
  - 4.1.2 Experiment and Task Configuration . . . 16
  - 4.1.3 Data Collection and Platform Functionality . . . 16
  - 4.1.4 Future Capabilities . . . 16
- 4.2 System Behavior Model . . . 17
- 4.3 User Interface Specification and Requirements Refinement . . . 18
- 4.4 Conceptual Data Model . . . 21
  - 4.4.1 Users and Profiles . . . 21

4.4.2 Timelines . . . 21
4.4.3 Tasks and Versioning . . . 22
4.4.4 Experiments . . . 23
4.4.5 Trainings . . . 23
4.4.6 Comprehensive Schema . . . 24
4.5 Chapter Summary . . . 24

5 Implementation **26**
5.1 High-Level Component Architecture . . . 26
5.2 Backend Implementation . . . 26
5.2.1 Data Persistence and ORM . . . 27
5.2.2 API Layer and Data Validation . . . 27
5.2.3 Security and Testing . . . 28
5.3 Frontend Implementation . . . 29
5.3.1 State Management . . . 29
5.3.2 Implementation of the Timeline Editor . . . 29
5.3.3 Implementation of the Step Configuration and Execution . . . 30
5.3.4 The Experimental Engine: Implementation of the Timeline Model . . . 30
5.3.5 User Interface Implementation . . . 31
5.3.6 Extensibility and Future Work . . . 31
5.4 User Interface Overview . . . 32
5.4.1 Researcher Interface and Workflow . . . 32
5.4.2 Visual Experiment Design: The Timeline Editor . . . 33
5.4.3 Building Blocks: The Timeline Step Editor . . . 34
5.4.4 Participant Management and Data Export . . . 37
5.4.5 Participant Experience and Workflow . . . 38
5.5 Chapter Summary . . . 39

6 Experimental Study: Procedure and Results **41**
6.1 Usability and Feasibility Study with Children . . . 41
6.1.1 Participants and Setup . . . 41
6.1.2 Materials: The Test Experiment . . . 42
6.1.3 Procedure and Data Collection . . . 42
6.2 Results . . . 44
6.2.1 Platform Feasibility and Data Integrity . . . 44
6.2.2 Qualitative Usability Findings . . . 44
6.2.3 Quantitative Performance Data from Child Participants . . . 45
6.3 Discussion . . . 46
6.3.1 Validation of Platform Objectives . . . 47
6.3.2 Addressing the Research Question: How can a web platform effectively integrate assessment, training, and monitoring of executive functions in children aged 5 to 9? . . . 47
6.3.3 Implications and Limitations . . . 48
6.3.4 Future Work . . . 48
6.4 Chapter Summary . . . 49

7 Conclusion and Future Work **50**
7.1 Conclusion and Main Contributions 50
7.2 Limitations 51
7.3 Future Work 51

References **53**

A Summary of Reviewed Executive Function Tasks **56**

B Executive Function Tasks Categorized by EF Component **70**

List of Figures

2.1 PRISMA Flow Diagram detailing the study selection process, based on the PRISMA-ScR guidelines. 6

4.1 Use Case diagram showing the core interactions between users and the CAPE-KIDS platform. Researchers can create tasks, experiments, and training protocols, manage participants, and export collected results. Participants can register and perform the activities assigned to them. These actions reflect the platform's goal of supporting both scientific data collection and cognitive training. 17
4.2 Conceptual mockup of the interface for creating a new reusable task template. . . 18
4.3 Specification for the visual Timeline Editor, illustrating the block-based design for creating experimental flows. 19
4.4 Mockup of the configuration modal for a block of stimuli, specifying global and individual trial settings. 20
4.5 Interactive prototype created in Figma, showing the complete user journey from task creation to task execution. This prototype was used for advanced feedback and validation with domain experts. 20
4.6 Conceptual data model for the User and UserProfile entities. 21
4.7 Conceptual data model for the timeline structure. 22
4.8 Conceptual data model for the task versioning strategy. 22
4.9 Data model for the Experiment module and its related entities. 23
4.10 Data model for the Training module, illustrating its parallel structure to the Experiment module and its reliance on shared core entities. 24
4.11 Complete conceptual data model of the CAPE-KIDS platform. The diagram illustrates the main entities and their relationships across the platform's core modules: user management, experiment and training workflows, task configuration, timeline structure, and result collection. Key elements include the separation of users and profiles, reusable task templates with versioning, flexible timelines with conditional logic, and support for both experimental and training sessions with result tracking. 25

5.1 High-level component architecture of the CAPE-KIDS platform. The diagram illustrates how the main technological components are deployed and interact in production. The frontend (a Next.js SPA) is hosted on Vercel and communicates with a backend REST API (NestJS) deployed on Railway. The backend handles business logic, data persistence via a PostgreSQL database, and media storage operations via Cloudflare R2. 27
5.2 Example of Zod schemas implemented in the shared library to ensure type-safe validation of the `metadata` field for different user profiles. 28

5.3 Swagger UI documentation for an API endpoint, where the example schema is automatically generated from the corresponding Zod schema. 29
5.4 The platform's login screen, providing secure access and language selection. . . . 33
5.5 The main management interface for Experiments, demonstrating the list view and the contextual action menu for a study. 33
5.6 The main Timeline Editor, showing a simple, linear flow for an experiment. . . . 34
5.7 The expanded view of a "Go/No-Go" task template, showing its own internal, self-contained timeline. 35
5.8 The configuration view for a `Custom Block`, showing the 16:9 screen editor with text, images, and a keyboard trigger defined. 35
5.9 The configuration modal for a `Sequential Stimuli` block, showing block-level settings and a list of individual stimuli. 36
5.10 The editor for an individual stimulus, where its visual content and specific response trigger are configured. 36
5.11 The interface for adding participants to a study from a list of available users on the platform. 37
5.12 The participant management dashboard for a specific experiment. The open action menu illustrates the options to download the collected data for a completed session. 38
5.13 The participant's simplified dashboard, showing a list of enrolled experiments and the action to start a session. 39
5.14 An example of an instruction screen presented to the participant before a task begins. 39
5.15 The notification message displayed to a participant attempting to re-access a completed experiment. 40

6.1 Instruction screen presented to children during the Flanker task (Portuguese version). Session monitored by a researcher specialist. 43

List of Tables

6.1 Summary of Quantitative Performance Data from Child Participants 46

A.1 Summary of Reviewed Executive Function Tasks 57

B.1 Executive Function Tasks Categorized by EF Component 70

Abbreviations and Symbols

API	Application Programming Interface
DCCS	Dimensional Change Card Sort
EFs	Executive Functions
ISI	Inter-Stimulus Interval
JWT	JSON Web Token
ORM	Object-Relational Mapper
SPA	Single-Page Application
WCST	Wisconsin Card Sorting Task

Chapter 1

Introduction

1.1 Context and Relevance

Executive functions (EFs) are a set of top-down cognitive processes that enable conscious control over thought and action to achieve goals [5]. Its core components include **inhibitory control**, **working memory**, and **cognitive flexibility** [5, 19]. These functions are fundamental during childhood, providing the foundation for key developmental milestones such as academic success, emotional regulation, and social adaptation [3]. The period between the ages of 5 and 9 is particularly critical, as it marks a phase of rapid development for these abilities [3].

1.2 Problem Statement

Despite the recognized importance of EFs, the tools and methodologies used for their assessment and training in children face significant challenges. A systematic scoping review of the scientific literature, conducted as part of this research, identified several key limitations in existing approaches:

- **Lack of Standardization:** Widely used assessment tasks often exhibit considerable variations in their design and scoring, making it difficult to compare findings across studies [6].
- **Limited Accessibility and Usability:** Many digital tools require programming skills, creating a technical barrier for researchers and psychologists. Furthermore, many traditional methods lack the ecological validity to capture performance in real-world contexts [6, 22].
- **Underutilization of Integrated Approaches:** There is a significant opportunity for platforms that not only assess EFs but also integrate training and monitoring protocols. Innovative approaches, such as the inclusion of physical activity, are also rarely implemented [26].

1.3 Proposed Solution: The CAPE-KIDS Platform

To address these gaps, this dissertation proposes the specification and implementation of **CAPE-KIDS (Cognitive Assessment and Personalized Experimentation for Kids)**, a web-based platform designed to provide an integrated environment for the assessment, training, and monitoring of EFs in children.

The conceptual design of CAPE-KIDS is centered on providing a unified and scientifically robust tool that is also highly accessible to non-technical users. Its distinguishing features, which form the basis of the system's specification, include:

- A **modular architecture** that separates concerns and allows for future expansion.
- A **visual, timeline-based interface**, which allows researchers to design complex and adaptive protocols without requiring programming skills.
- A system of **configurable and versioned task templates** derived from the literature, promoting methodological standardization and reproducibility.

1.4 Research Question and Objectives

Based on the identified problem, the development of this work was guided by the following primary research question:

> *How can a web-based platform be effectively designed to integrate assessment, training, and monitoring of executive functions in children aged 5-9?*

To answer this question, the following objectives were established for the project:

1. To analyze the current state of the art by conducting a systematic literature review to map the most commonly used executive function tasks and their implementation strategies.
2. To define the specific requirements for the CAPE-KIDS platform through a series of meetings with a domain expert in developmental neuroscience from the University of Minho.
3. To design and implement a functional prototype based on the specified requirements, focusing on a modular architecture and a no-code interface for researchers.
4. To conduct an initial evaluation of the prototype's usability and feasibility through a pilot implementation with child participants.

1.5 Dissertation Structure

This document is organized into seven chapters. Following this introduction, **Chapter 2** presents a review of the state of the art on executive functions and their assessment. **Chapter 3** details the overall research methodology for the project.

The core technical contributions are presented in the subsequent chapters. **Chapter 4** is dedicated to the **System Specification**, presenting the technology-agnostic design of the CAPE-KIDS platform. **Chapter 5** then details the **Implementation** of this specification, describing the concrete technologies and engineering decisions made.

Chapter 6 presents the complete experimental validation of the platform. It details the study's procedure, presents the collected quantitative and qualitative results, and discusses their implications in response to the research question. Finally, **Chapter 7** provides the overall conclusion of this work, acknowledges its limitations, and proposes directions for future work.

Chapter 2

State of the Art

This chapter establishes the scientific and technical context that underpins this dissertation. It begins by providing an overview of the foundational literature on Executive Functions (EFs), their core components, and their critical role in child development. Following this, the chapter transitions from a general overview to a detailed and systematic analysis of the current landscape of digital tools and tasks used for EF assessment.

To ground this project in empirical evidence and move beyond a simple literature review, a formal **systematic scoping review** was conducted. This chapter details the methodology of this review, presents its key findings, and discusses their implications. The analysis identifies significant, persistent gaps in the field related to standardization, accessibility, and the integration of training and monitoring functionalities. These identified challenges form the core scientific rationale for the research question and the subsequent design of the CAPE-KIDS platform.

2.1 Executive Functions: Foundational Concepts

Executive Functions are a family of top-down mental processes required for goal-directed behavior, particularly in situations that are novel or require concentration [5]. While various models exist, there is a broad consensus around three core, interrelated EFs that provide the foundation for higher-order cognitive skills like reasoning and problem-solving [19]:

Inhibitory Control is the ability to deliberately suppress dominant, automatic, or prepotent responses. This includes both response inhibition (self-control) and interference control (the ability to ignore distracting information) [5, 19].

Working Memory refers to the capacity to hold information in mind and actively manipulate it. It involves not just temporary storage but also active manipulation of information, which is critical for tasks like following multi-step instructions or performing mental arithmetic [5, 3].

Cognitive Flexibility (also known as shifting) is the ability to adapt one's thinking and behavior in response to changing demands or rules. It allows for creative problem-solving and seeing situations from different perspectives [5, 19].

Childhood represents a critical period for EF development, with significant growth occurring between the ages of 5 and 9 years [3]. Assessing and strengthening these functions in school-aged children is essential for promoting academic achievement and reducing educational disparities. Research has consistently demonstrated that EF skills correlate with mathematics, reading comprehension, and writing proficiency, sometimes even more strongly than IQ [25].

2.2 The State of Digital EF Assessment

While traditional methods like paper-based tests and rating scales exist, they often lack real-world applicability and cultural adaptability [6]. This has led to a growing interest in digital platforms that allow for real-time, research-oriented EF assessment. However, the digital landscape itself presents challenges, including a lack of clarity in defining task parameters and a limited integration of innovative approaches like physical activity [26].

To effectively address the research question of this dissertation, a foundational investigation into the current state of the art was required. Therefore, a formal scoping review was conducted to analyze the existing landscape. This review was guided by the following objectives:

1. To map the types of executive function tasks most commonly used in recent literature for children aged 5-9.

2. To analyze the main implementation strategies and metrics used to assess performance in these tasks.

3. To identify the persistent challenges and limitations of current approaches, thereby establishing the requirements for a new solution.

2.3 Review Methodology

The review was structured according to the PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) guidelines to ensure a transparent and reproducible method. A systematic search was performed in the SCOPUS, ScienceDirect, and ERIC databases for studies published between 2019 and 2024, this time frame was chosen to ensure that the review captures the most recent developments in executive function (EF) assessment methods while maintaining a manageable scope, allowing for a more feasible data collection process, and still ensuring the inclusion of studies reflecting current research trends. The search query was: `"executive function task" AND (children OR child OR childhood)`.

Studies were selected based on strict inclusion criteria, including a focus on typically developing children aged 5 to 9 years. After a multi-stage screening process detailed in Figure 2.1, which

involved removing duplicates and assessing 113 full-text reports for eligibility, a final selection of **23 studies** was included for analysis.

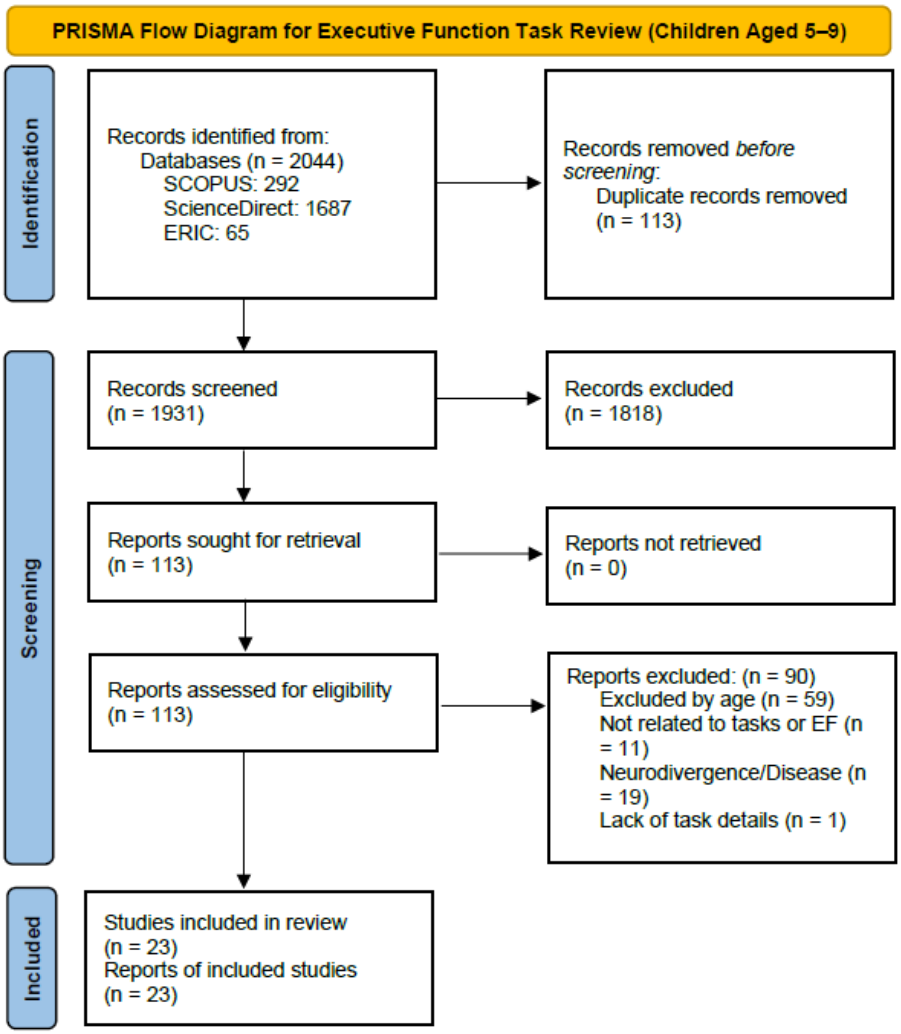

Figure 2.1: PRISMA Flow Diagram detailing the study selection process, based on the PRISMA-ScR guidelines.

2.4 Key Findings

The findings from this review highlight both the progress and persistent challenges in the assessment of executive functions (EF) in children. While a variety of tasks have been developed to measure inhibition, working memory, and cognitive flexibility, substantial inconsistencies exist in their **implementation, trial parameters, and scoring methods**. These discrepancies pose significant barriers to the **comparability of results across studies**, ultimately limiting the field's ability to establish robust conclusions regarding EF development.

To better illustrate the landscape of EF assessments, Table B.1 provides a structured classification of tasks according to their primary EF component—Inhibitory Control, Working Memory, or

Cognitive Flexibility. This categorization highlights the diversity of tasks used in research and reveals an uneven distribution of EF components, with some functions receiving significantly more focus than others. Furthermore, Table A.1 presents a comprehensive synthesis of EF tasks, detailing their methodologies, implementation strategies, and scoring mechanisms, reinforcing the methodological inconsistencies across studies. These insights emphasize the need for a more standardized and structured approach to EF task design and implementation.

2.4.1 Lack of Standardization Across Studies

One of the main issues identified is the **lack of consistency in task parameters**, even for widely used assessments. Tasks such as the *Flanker Task*, *Go/No-Go Task*, and *Dimensional Change Card Sort (DCCS)* exhibit considerable **variations in trial lengths, block structures, and response criteria**, which directly impact the difficulty and interpretation of results. For example, in the *Go/No-Go Task*, the proportion of "go" versus "no-go" trials varies across studies, affecting inhibitory demands and making comparisons difficult. Similarly, in the *Flanker Task*, discrepancies in stimulus presentation times and response deadlines lead to inconsistent difficulty levels across implementations.

Beyond these technical aspects, scoring criteria also differ widely. Some tasks, particularly those embedded in standardized cognitive batteries (e.g., *Backward Digit Span*), benefit from well-established norms, allowing results to be contextualized within broader developmental frameworks. However, many other EF tasks lack reference data, making it unclear **how performance should be interpreted across different samples and cultural settings**. Without standardized guidelines, results remain difficult to compare, limiting the generalizability of EF research.

2.4.2 The Need for a Unified Framework

Given these challenges, a structured and **standardized digital platform** is necessary to enhance the comparability and reliability of EF assessments. Such a platform would serve to:

- **Standardize task implementation** – Define fixed parameters for trial numbers, timing, and scoring to ensure consistency across studies.
- **Ensure cross-cultural applicability** – Adapt tasks to different linguistic and educational contexts while maintaining their core cognitive demands.
- **Facilitate large-scale data collection** – Allow researchers to gather and compare EF performance across diverse populations with greater methodological rigor.
- **Integrate physical activity into digital assessments** – Many EF tasks remain **static**, despite evidence that **motor engagement enhances cognitive performance** [26]. A modern framework should explore **ways to incorporate movement-based tasks digitally**, using motion tracking, adaptive feedback, or hybrid digital-physical designs.

Tasks that are highly structured and **less dependent on environmental context**, such as the *Flanker Task*, *DCCS*, and *Go/No-Go Task*, are well-suited for digital adaptation and could serve as the foundation for a standardized framework. In contrast, **movement-based EF tasks**, such as *HTKS* and *Freeze Task*, require innovative **digital adaptations** to preserve their cognitive and motor demands while ensuring consistency across studies.

Additionally, some tasks present specific characteristics that make them particularly promising for digital adaptation. For **inhibitory control**, tasks such as the *Stroop Task*, *Flanker Task*, and *Go/No-Go Task* have well-established methodologies and are widely implemented in computerized settings. In the case of **working memory**, tasks like the *Backward Digit Span*, *Corsi Blocks Task*, and *Keep Track Task* provide reliable cognitive measures and can be effectively implemented in digital formats with minimal modifications. Finally, for **cognitive flexibility**, the *Dimensional Change Card Sort (DCCS)*, *Wisconsin Card Sorting Task (WCST)*, and *Dots Task* offer strong candidates for digital transformation due to their structured nature and adaptability to automated scoring.

2.4.3 Main Executive Function Tasks Identified

Inhibitory Control Tasks The review found that tasks measuring inhibition were the most common. The **Flanker Task** and the **Stroop Task** were frequently used to measure interference control, requiring participants to focus on a target stimulus while inhibiting distracting information. For response inhibition, the **Go/No-Go Task** was the dominant paradigm, requiring participants to respond to a frequent stimulus but withhold that response for a rare one.

Working Memory Tasks For working memory, tasks that required both storage and manipulation of information were prevalent. **Backward Span Tasks** (e.g., Backward Digit Span, Corsi Blocks) were very common, as they require participants to recall a sequence in reverse order, thus engaging the active manipulation component of working memory.

Cognitive Flexibility Tasks To assess the ability to shift between mental sets, the **Dimensional Change Card Sort (DCCS)** was the most widely used task. In this task, children must switch between sorting cards by one rule (e.g., color) to another (e.g., shape), providing a clear measure of cognitive flexibility.

2.5 Challenges and Opportunities

Beyond mapping the tasks, the review identified several critical challenges in the current state of the art.

The Challenge of Standardization The most significant finding is a widespread **lack of standardization**. The review revealed that even for classic paradigms like the Flanker Task, the implementation details—such as the number and timing of trials, the specific stimuli used (e.g., arrows

vs. fish), and the scoring criteria—vary substantially across studies. This inconsistency makes it extremely difficult to compare or replicate findings, undermining the cumulative nature of the science. This points to an urgent need for a framework that can support the creation and distribution of standardized task implementations.

The Need for an Integrated and Accessible Framework The diversity of the 38 tasks identified, each with unique parameters and configurations, highlights a significant implicit challenge: the technical effort required for a researcher to implement these tasks from scratch for each new study. The lack of a unified tool that provides access to a wide range of pre-configured, standardized tasks creates a de facto **accessibility barrier**. Furthermore, the review revealed that research protocols often combine assessment and training. The absence of platforms that seamlessly **integrate assessment, training, and monitoring** functionalities forces researchers to use multiple, disconnected tools, complicating their workflow.

Opportunities for Integrated and Innovative Design The review revealed two significant missed opportunities. First, there is a clear need for platforms that can **integrate assessment, training, and monitoring** within a single ecosystem, as most tools focus on only one of these aspects. Second, despite strong evidence that physical activity enhances cognitive performance in children [26], its **minimal integration** into EF tasks remains a largely untapped area for innovation.

2.6 Limitations

This review has some limitations that should be acknowledged.

First, the inclusion criterion focusing on children aged 5 to 9 years may have excluded studies with tasks designed for younger children, especially 4-year-olds. While this age range was strategically chosen to align with the transition to formal schooling and the platform's intended applicability, it may limit the generalizability of findings to earlier developmental stages. Tasks for younger children could offer additional insights into early EF development and inform future adaptations of the platform for broader age groups.

Second, the review included only studies published between 2019 and 2024. This decision helped to focus on recent and relevant research but may have left out older studies that still offer useful methods or theoretical insights.

Finally, this review did not include a formal evaluation of existing digital platforms. The focus was on analyzing executive function tasks and how they are used in recent research, rather than comparing the platforms that implement them. However, an informal review of widely used tools, such as E-Prime and Gorilla, was conducted to better understand current limitations in terms of flexibility, accessibility, and integration of assessment and training. A more detailed comparison of platforms could be included in future work.

2.7 Conclusion and Implications for Platform Design

This review systematically mapped the executive function tasks most commonly used with children aged 5 to 9, identifying patterns in their implementation, metrics, and cognitive targets. While it did not directly evaluate digital platforms, the analysis uncovered recurring challenges—chiefly the lack of standardization, the fragmented nature of EF assessment and training protocols, and the underuse of adaptive or integrative features.

These findings provided a strong foundation for defining the functional and conceptual requirements of a new tool. The CAPE-KIDS platform was conceived not as a generic assessment engine, but as a direct response to these gaps—offering a unified environment where scientifically grounded EF tasks can be configured, deployed, and monitored in a reproducible and flexible manner.

While platforms such as Cambridge Cognition's CANTAB and Gorilla provide robust tools for executive function assessment and experimentation, they are often constrained in different ways—CANTAB by its fixed task repertoire, and Gorilla, while highly flexible, by the lack of specific features tailored to child-friendly design or longitudinal developmental studies. CAPE-KIDS, in contrast, was designed to be open, customizable, and accessible to researchers, supporting configurable EF **assessment**, **training**, and **longitudinal monitoring** within a single unified environment.

Moreover, the platform was designed to accommodate future integrations with **wearables**, **neuroimaging markers**, and other kinds of triggers, making it extensible to other age groups or cognitive domains. The customizable configuration of tasks enables researchers to adjust stimuli, parameters, and task flow without compromising scientific rigor, ensuring both experimental control and ecological validity.

By translating insights from the literature into concrete software architecture and design choices, this dissertation bridges the current methodological fragmentation and aims to empower researchers and educators with an accessible, evidence-informed platform. The following chapters present how these conceptual insights were formalized into system specifications and implemented in a working prototype.

Despite advancements in digital tools for EF assessment, there remains a gap in platforms specifically designed to support customizable, research-oriented workflows tailored to school-aged children. Existing solutions often lack support for adaptive training protocols, integration of multimodal data (e.g., wearables, neuroimaging), or flexible experimental control usable by non-technical researchers. Given that EF skills are strongly linked to academic performance in areas such as mathematics, reading, and writing—sometimes even more so than IQ—there is a pressing need for accessible, evidence-based tools that allow both assessment and longitudinal monitoring. In this context, the CAPE-KIDS platform emerges as a response to these challenges, aiming to fill the methodological and practical gaps identified in the literature.

2.8 Chapter Summary

In summary, this chapter defined the core concepts of executive functions and presented a detailed analysis of the state of the art based on a systematic scoping review. This review process confirmed that while a diverse set of digital tasks exists, the field is significantly hampered by challenges in standardization, accessibility, and integration. These identified gaps establish a clear and urgent need for a new generation of tools. The following chapters will detail the specification and implementation of the CAPE-KIDS platform, a solution engineered specifically to address these challenges.

Chapter 3

Research Methodology

This chapter details the methodological framework that guided this dissertation from its inception to its final validation. Following the identification of key challenges in the scientific literature (as detailed in Chapter 2), a phased research approach was adopted. This chapter outlines this overall approach and then details the specific methodologies used for the platform's development and its multi-part evaluation process.

3.1 Overall Research Approach

The research was structured in three main phases, moving from problem discovery to solution implementation and evaluation.

1. **Phase 1: Problem Identification and Requirements Specification.** It involved a two-pronged approach to establish the project's requirements. First, as described in Chapter 2, a formal scoping review was conducted to analyze the state of the art and identify general gaps in existing tools. Second, these findings were complemented by a series of meetings with a domain expert in child psychology to define the specific requirements for the CAPE-KIDS platform.

2. **Phase 2: Solution Design and Implementation.** Based on the requirements from Phase 1, a functional prototype of the CAPE-KIDS platform was designed and implemented, with its detailed specification and implementation described in Chapters 4 and 5, respectively. The goal was to create a high-fidelity proof of concept demonstrating the feasibility of the proposed architecture.

3. **Phase 3: Evaluation and Validation.** The final phase focuses on evaluating the developed prototype through a two-part process. The first part involves a *formative evaluation* with subject-matter experts. The second part, a *summative* pilot study with end-users (children), was then conducted to assess the platform's usability in a practical setting.

3.2 Platform Development Methodology

The development of the CAPE-KIDS prototype followed an iterative process, guided by a set of core principles established at the outset of the project. The primary goal was to create a scientifically valid, extensible, and usable tool for researchers without a technical background. The platform's design was based on a modular, timeline-based architecture conceived to meet these principles. The development process was also directly informed by the formative evaluation with experts, allowing for design refinements based on their feedback.

3.3 Validation Methodology

To assess the viability and effectiveness of the CAPE-KIDS prototype, a two-part validation methodology was executed.

3.3.1 Formative Evaluation with Experts

A formative evaluation was conducted to validate the platform's core concepts and design from a researcher's perspective before end-user testing.

- **Participants:** The participants were two experts from the University of Minho, selected for their extensive research experience and publications in psychological assessment, executive functions, and developmental psychology. One expert holds a PhD in Psychology and is an Assistant Professor, while the other is a researcher with a PhD in Psychology specializing in neuropsychological assessment.
- **Procedure:** The evaluation consisted of structured interview sessions, each with a duration between 35 and 40 minutes. In each session, the project's objectives and main functionalities were presented, followed by an interactive, screen-by-screen demonstration of the system's high-fidelity prototype created in Figma. A "think-aloud" protocol was encouraged, where experts provided continuous verbal feedback while viewing the prototype.
- **Data Collection and Analysis:** All sessions were recorded, transcribed, and analyzed to extract key insights and recommendations. The qualitative data was analyzed thematically, with feedback categorized into themes such as Usability and User Experience, Scientific Validity, and Data Analysis. The results of this evaluation, which directly informed and validated the final design, are presented in detail in Chapter 4.

3.3.2 Methodology for the Experimental Study with Children

Following the formative evaluation, a pilot study was designed and conducted with the platform's target end-users to assess its real-world usability and feasibility.

- **Objectives:** To evaluate the usability of the participant-facing interface and to verify that the platform can capture performance metrics during EF tasks.

- **Participants:** The study involved a sample of four children aged 5 to 9 years.
- **Materials:** The materials consisted of the CAPE-KIDS platform running on a desktop device, with an experiment timeline containing a sequence of representative EF tasks.
- **Procedure:** The procedure involved each child participating in a single session, monitored by a researcher and a psychology specialist. After receiving standardized instructions, the children interacted with the experiment using a keyboard and mouse.
- **Data Collection:** The platform was used to automatically record quantitative, trial-by-trial data, which was complemented by qualitative observational notes taken by the researcher.

The detailed execution and results of this study is described in Chapter 6.

3.4 Chapter Summary

This chapter described the research methodology used throughout this dissertation. The work followed a three-phase structure: identifying the problem and requirements, designing and implementing the solution, and validating the platform. The development approach was iterative and user-centered, guided by expert feedback and focused on creating an accessible, extensible tool. The validation process included a formative evaluation with experts and a pilot study with children, confirming both the platform's scientific alignment and practical feasibility. This methodological framework provided a solid foundation for the CAPE-KIDS platform and supported the main goals of the project.

Chapter 4

System Specification

This chapter details the specification for the CAPE-KIDS platform. Based on the requirements identified in the previous chapters, it defines the system's expected functionalities, its conceptual data structure, its behavioral models, and its user interface design. This specification serves as the formal blueprint for the concrete implementation detailed in Chapter 5.

4.1 Requirements

The following requirements, which are grouped by their functional area, were defined in collaboration with a domain expert in child psychology. The initial set of requirements was gathered based on initial meetings. This set was then iteratively refined and validated during meeting sessions, where mockups of the platform were presented to the expert, allowing to receive feedback to directly shape the final specification. To better organize the development process and define the scope for the initial prototype, the requirements were prioritized using the MoSCoW method, classifying each as **MUST** (essential for the core functionality), **SHOULD** (important, but not critical for the initial release), or **COULD** (desirable, but could be postponed for future versions).

4.1.1 User and Participant Management

[MUST] R1 The system must allow researchers to register on the platform and subsequently enroll participants in their studies.

[MUST] R2 The system must securely store essential participant information (e.g., age, gender) and allow for extensible metadata for different profile types.

[MUST] R12 The system must implement a robust authentication mechanism to provide secure access for registered users.

[SHOULD] R17 The system should allow a study owner to invite and manage other collaborating researchers, defining their permission levels (e.g., view, edit).

4.1.2 Experiment and Task Configuration

[MUST] R3 The system must enable researchers to configure different executive function tasks based on a set of core templates.

[MUST] R4 The configuration must provide options to customize key task parameters, such as stimulus duration, number of trials, and inter-stimulus intervals.

[MUST] R11 The system must provide a web platform for all task and experiment configuration.

[MUST] R13 The system must support the inclusion of visual and/or auditory instructions to guide participants.

[MUST] R14 The system must support the inclusion of a demonstration trial to familiarize participants with the task mechanics.

[MUST] R15 The system must support the implementation of practice trials under controlled conditions.

[SHOULD] R16 The design should support adaptive difficulty levels, where task parameters can adjust based on participant performance.

[SHOULD] R18 The design should allow the scientist to create training sessions for development of the executive functions.

4.1.3 Data Collection and Platform Functionality

[MUST] R5 The system must record participant response data in real time, including the response itself and its latency.

[MUST] R6 Data must be stored in a centralized, persistent data store on a server.

[MUST] R9 The system must be able to display immediate visual and/or auditory feedback to the participant.

[SHOULD] R10 The platform should provide basic performance reports or data export capabilities for researchers.

4.1.4 Future Capabilities

[COULD] R7 The system could be extended to detect specific movements (such as jumping or crouching) using a device's accelerometer and gyroscope.

[COULD] R8 The architecture could support future integration with wearable devices for more precise data collection.

4.2 System Behavior Model

User interactions with the CAPE-KIDS platform can be summarized through a set of high-level use cases that reflect the system's main functionalities. These use cases are directly based on the functional requirements described earlier and help ensure that the system supports the key needs of both researchers and participants.

Figure 4.1 shows the two main actors: the **Researcher**, who is responsible for creating studies and managing participants, and the **Participant**, who performs tasks as part of an experiment or training. Each actor interacts with the platform in different ways.

Researchers can:

- Register and log into the system securely.
- Configure experiments, cognitive tasks, and training sessions.
- Enroll and manage participants linked to their studies.
- Access and export performance results for analysis.

Participants can:

- Create an account and access the platform.
- Execute assigned experiments and training protocols, following instructions and performing tasks.

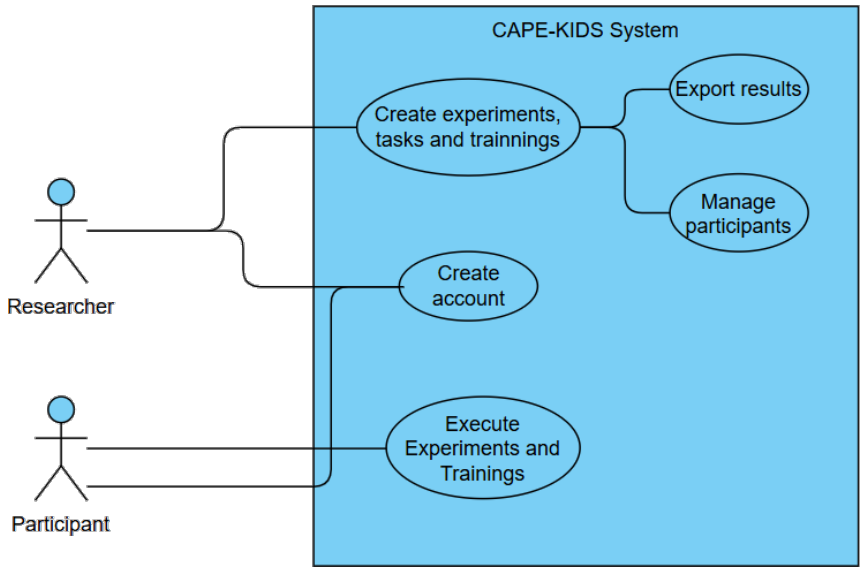

Figure 4.1: Use Case diagram showing the core interactions between users and the CAPE-KIDS platform. Researchers can create tasks, experiments, and training protocols, manage participants, and export collected results. Participants can register and perform the activities assigned to them. These actions reflect the platform's goal of supporting both scientific data collection and cognitive training.

4.3 User Interface Specification and Requirements Refinement

Following the initial definition of requirements, a series of conceptual mockups, or wireframes, were designed in Figma. These visual artifacts were used as an interactive tool during the requirements refinement sessions with the domain expert. By presenting and discussing these visual concepts, it was possible to gather immediate and concrete feedback, allowing the expert's insights to directly shape and validate the final set of functional requirements.

This section presents the key mockups that served as this specification and validation tool, detailing the specified workflow for a researcher.

The process begins with the creation of a reusable **Task Template**. The mockup in Figure 4.2 specifies the simple interface for a researcher to define a new task's name and description. This serves as the foundational block that can later be incorporated into multiple experiments or trainings.

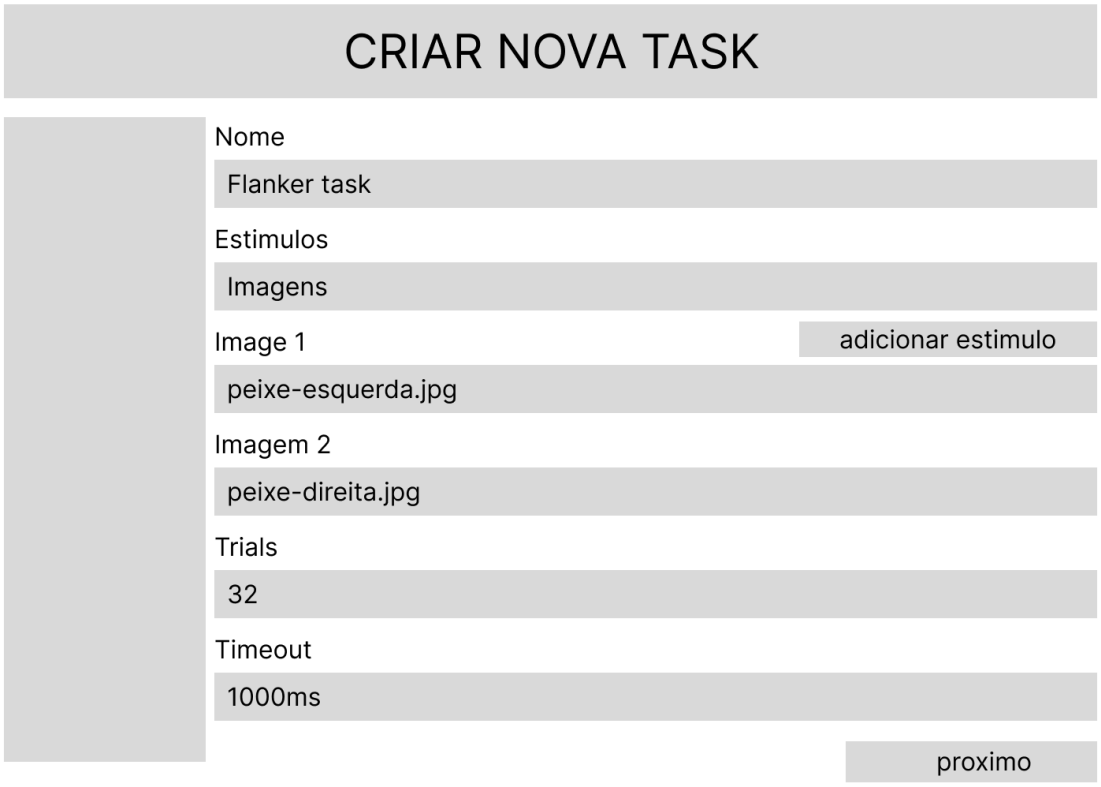

Figure 4.2: Conceptual mockup of the interface for creating a new reusable task template.

Once task templates are available, the core of the researcher's workflow is specified in the visual **Timeline Editor** (Figure 4.3). This mockup defines an interface where steps of an experiment can be visually arranged and connected. This design is central to the requirement of providing a no-code experience for creating complex experimental flows.

The specification requires that each step in the timeline be highly configurable. Figure 4.4 shows the mockup for the configuration modal of a stimuli block. It specifies the necessary fields



Figure 4.3: Specification for the visual Timeline Editor, illustrating the block-based design for creating experimental flows.

for a researcher to control block-level parameters (e.g., number of trials, timing) as well as manage a list of individual stimuli within that block.

By presenting and navigating through this initial wireframe, it was possible to discuss and validate the specified user journey, ensuring it was both logical and aligned with the project's requirements.

After validating the wireframes, a high-fidelity prototype was developed to simulate the real user experience more accurately. This prototype included interactive navigation flows and realistic UI elements (see Figure 4.5). It was used in further feedback sessions with domain experts, enabling a more detailed evaluation of usability, flow clarity, and overall system logic.

The flow implemented in the prototype was designed to reflect a typical use case of the platform. The user begins by creating a new cognitive task composed of two steps. In each step, specific visual stimuli are configured: the user adds images (a mouse and a tiger) and defines keyboard-based triggers — pressing the "D" key when the mouse appears on the left, and the "F" key when the tiger appears on the right. The user then adjusts task parameters such as the inter-stimulus interval, stimulus duration, and whether the presentation order should be randomized.

Following the configuration, the prototype presents a preview of the task as it would appear to a participant. This includes the full sequence of stimuli, instructions, and feedback screens, providing a clear representation of the task execution flow. At the end of the preview, an example

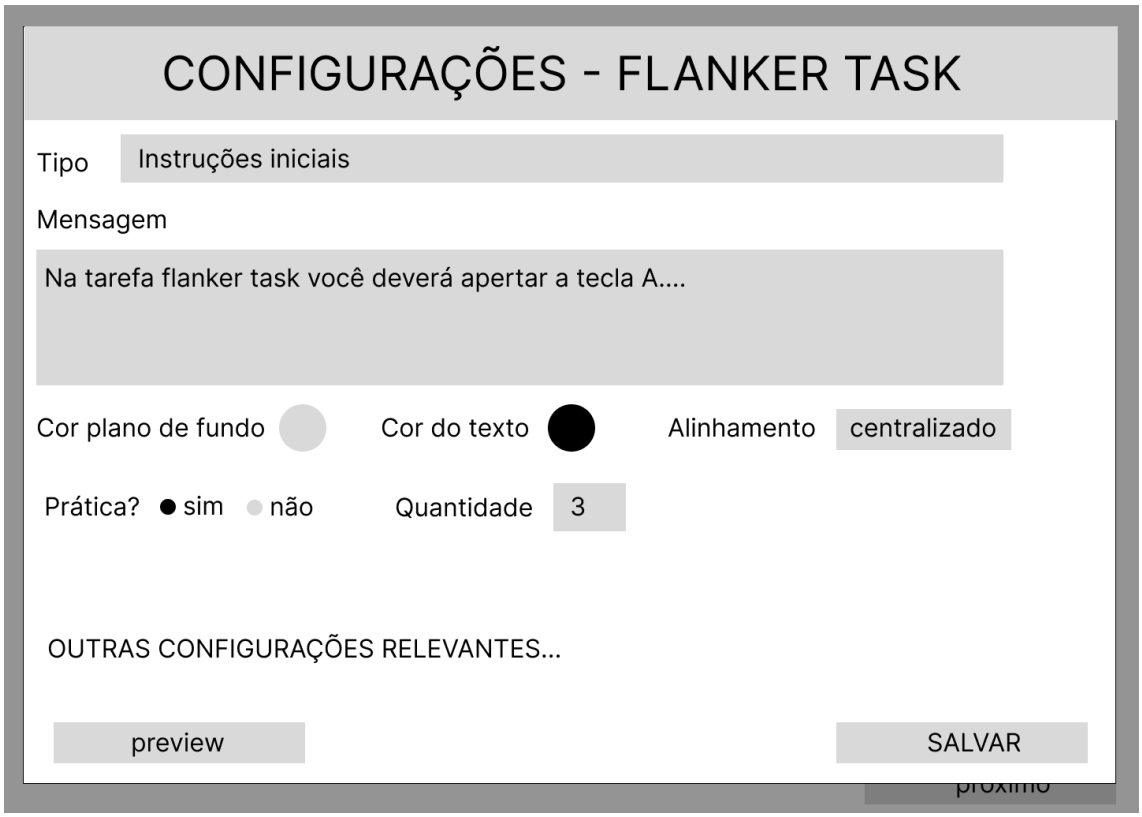

Figure 4.4: Mockup of the configuration modal for a block of stimuli, specifying global and individual trial settings.

of the data output is shown, illustrating how participant responses would be recorded and stored by the system.

This end-to-end simulation allowed domain experts to validate not only the interface, but also the logic and behavior of the system from the researcher's and participant's perspectives.

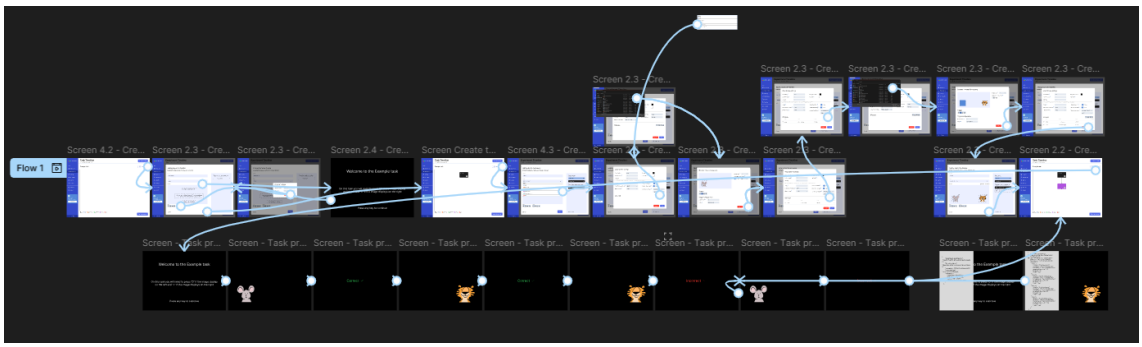

Figure 4.5: Interactive prototype created in Figma, showing the complete user journey from task creation to task execution. This prototype was used for advanced feedback and validation with domain experts.

The final set of requirements presented earlier in this chapter was refined and validated through this iterative design process, combining low- and high-fidelity mockups with real-time expert in-

put. This ensured that the system's functionality and user interface were both technically feasible and domain-appropriate.

4.4 Conceptual Data Model

The data model for the CAPE-KIDS platform was designed to support its dual core functionalities: structured assessment via **Experiments** and cognitive intervention via **Trainings**. To ensure both a solid structure and future extensibility, the conceptual schema combines strictly defined entities with flexible attributes.

For certain entities that require extensible information—such as user profiles—the model specifies a flexible metadata attribute. The structure of this attribute is not arbitrary; it must conform to a set of predefined validation schemas. This approach guarantees data integrity and consistency throughout the system, allowing for rich, structured information to be stored for different roles (e.g., a researcher versus a participant) while maintaining a coherent underlying model.

The following subsections detail the key logical entities of the data model, grouped by their functional domain.

4.4.1 Users and Profiles

The user management module separates authentication data (`User` entity) from profile-specific data (`UserProfile` entity) for security and modularity. This allows the `UserProfile` to contain a flexible metadata field whose structure is dependent on the user's role, such as storing institutional details for a researcher or demographic data for a participant. The conceptual model for these entities is illustrated in Figure 4.6.

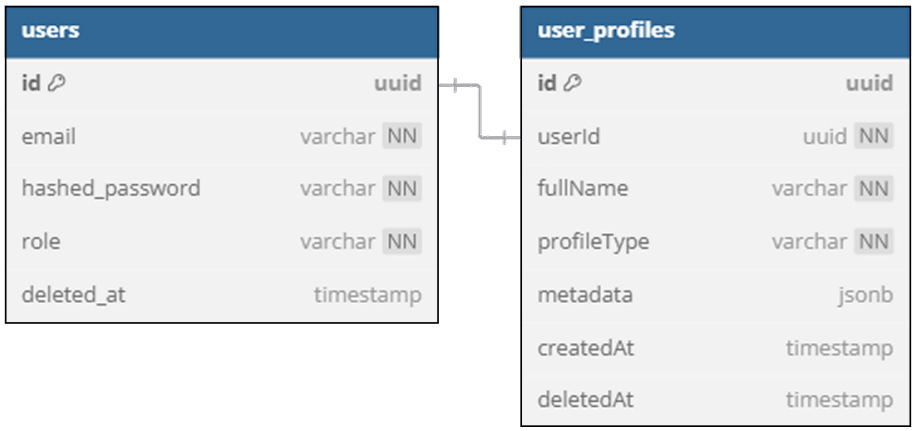

Figure 4.6: Conceptual data model for the User and UserProfile entities.

4.4.2 Timelines

The structure of any experiment, training, or task is defined by a `Timeline`. As shown in Figure 4.7, this core entity holds an ordered sequence of `TimelineStep` records. The connections and

branching logic between these steps are managed by the `StepConnection` entity, allowing for the creation of complex adaptive flows.

A key design choice within the `Timeline` entity is the use of a polymorphic association via the `sourceType` and `sourceId` columns. This allows a single timeline structure to be linked to different parent entities—such as an `Experiment` or a `Training`. This approach greatly enhances the system's modularity and future extensibility, as new features that require a sequential flow can reuse this core timeline module without any changes to the database schema.

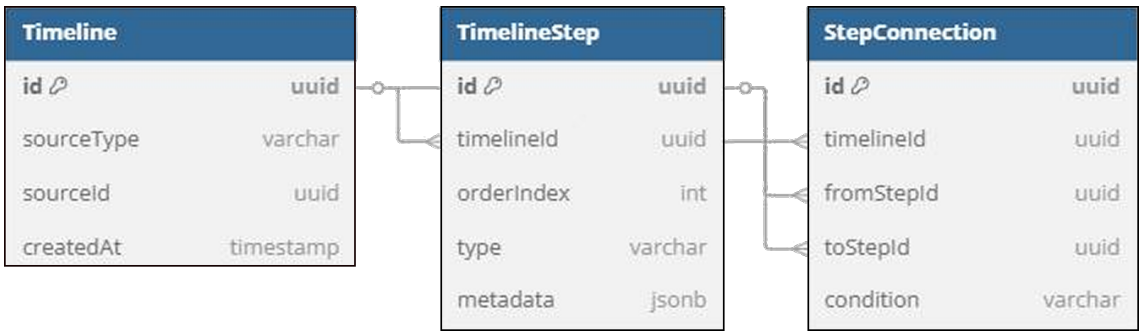

Figure 4.7: Conceptual data model for the timeline structure.

4.4.3 Tasks and Versioning

To ensure scientific reproducibility, the data model specifies a critical versioning strategy. It distinguishes between a `Task` (a reusable master template) and a `TaskVersion` (an immutable snapshot of that template's configuration). As illustrated in Figure 4.8, this immutable template pattern guarantees that once a study is created, its conditions remain fixed and reproducible.

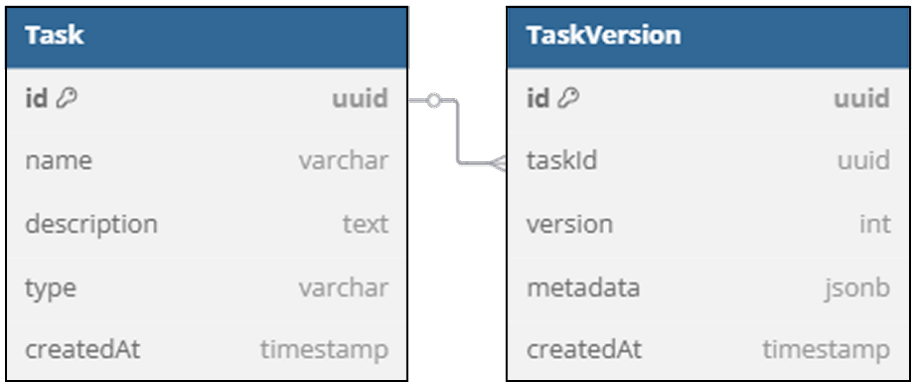

Figure 4.8: Conceptual data model for the task versioning strategy.

In practice, this versioning model ensures that each study uses a specific snapshot of a task's configuration. This design guarantees reproducibility and isolation, allowing different studies to use the same base task while preserving their own modifications independently.

4.4.4 Experiments

The `Experiment` entity serves as the container for a formal research study. It is linked to its creator, collaborating scientists (`ExperimentScientist`), and enrolled participants (`ExperimentParticipant`). All data generated is stored in the `ExperimentResult` table, ensuring every response is traceable to a specific participant and experiment, as detailed in Figure 4.9.

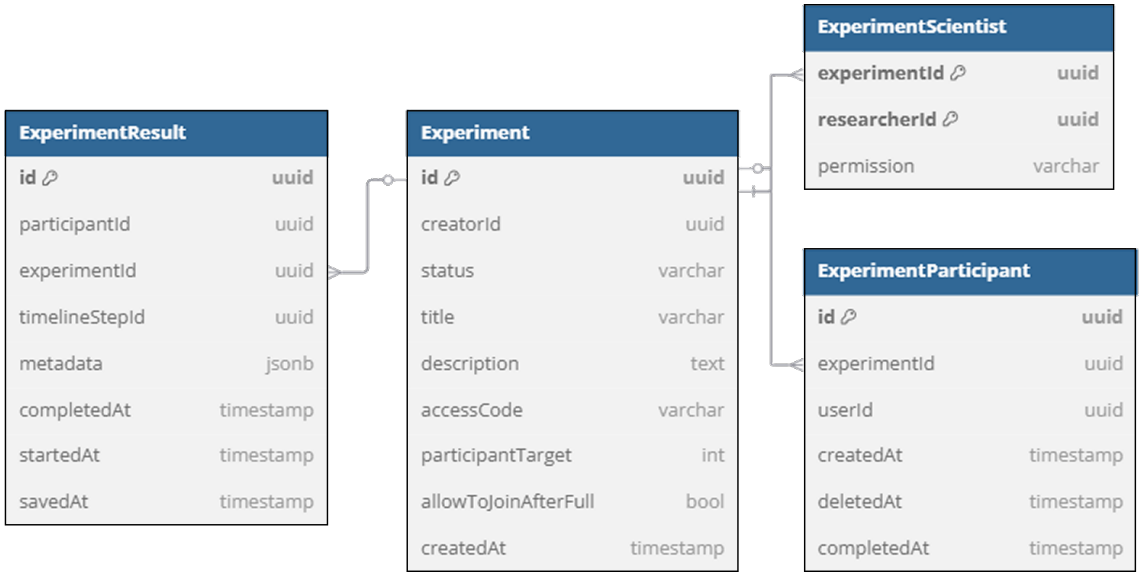

Figure 4.9: Data model for the Experiment module and its related entities.

To ensure scientific control, each experiment can only be executed once per participant. This constraint is a key specification of the data model, enforced through the logic associated with the ExperimentParticipant entity, which records a participant's completion of an experiment.

4.4.5 Trainings

Parallel to the experiment module, the `Training` entity is designed to support cognitive intervention protocols. This provides a distinct workflow for activities focused on practice and skill improvement rather than formal, single-session assessment.

The module intentionally mirrors the architecture of an experiment, with its own top-level `Training` entity and associated collaborators (`TrainingScientist`). However, it differs in how participation is managed. In contrast to an experiment, which is designed as a single-shot assessment for each participant, a training protocol is built for repetition. The data model reflects this by allowing a single user to have multiple `TrainingSession` records for the same `Training`. This design enables longitudinal interventions, and the total number of sessions can be controlled via the `sessions_limit` property on the parent training entity, which defines how many times a participant may repeat the protocol.

Crucially, as shown in the data model (Figure 4.10), the training module reuses the exact same core `Timeline` and `Task` entities. This modularity is a key strength of the platform, as it allows

a complex cognitive task to be designed once and then seamlessly deployed for either assessment or intervention.

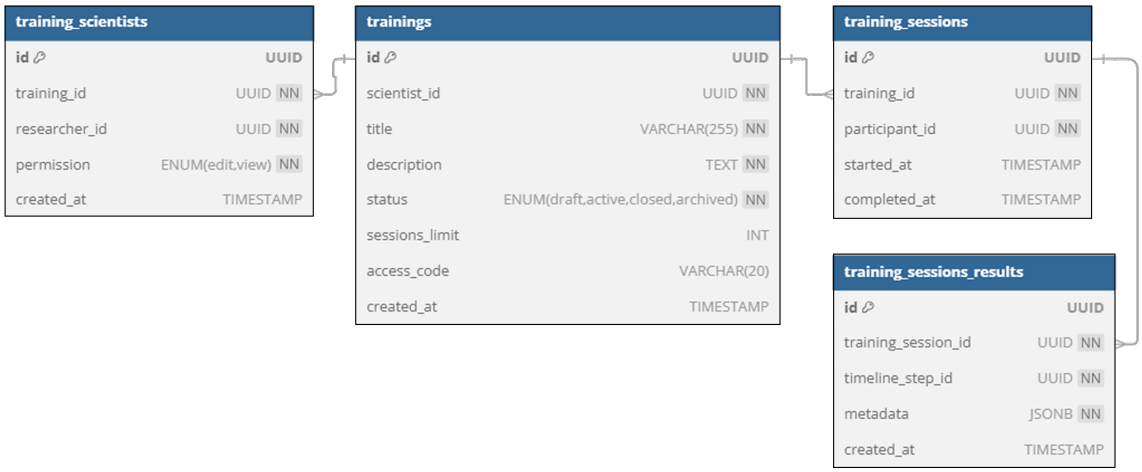

Figure 4.10: Data model for the Training module, illustrating its parallel structure to the Experiment module and its reliance on shared core entities.

4.4.6 Comprehensive Schema

Finally, Figure 4.11 presents the complete and integrated conceptual data model, illustrating how all the previously described modules and entities connect to form a cohesive system.

4.5 Chapter Summary

This chapter presented the formal specification of the CAPE-KIDS platform. It began by detailing the system requirements, organized using the MoSCoW prioritization method and validated through collaboration with a domain expert. The user interface was defined through mockups, enabling an iterative design process focused on usability and clarity. The conceptual data model was then described, emphasizing flexibility, modularity, and scientific rigor through elements like timeline structures and task versioning. Finally, the system's behavioral model was introduced using a use case diagram. Together, these elements formed a complete and well-structured blueprint for the platform's implementation.

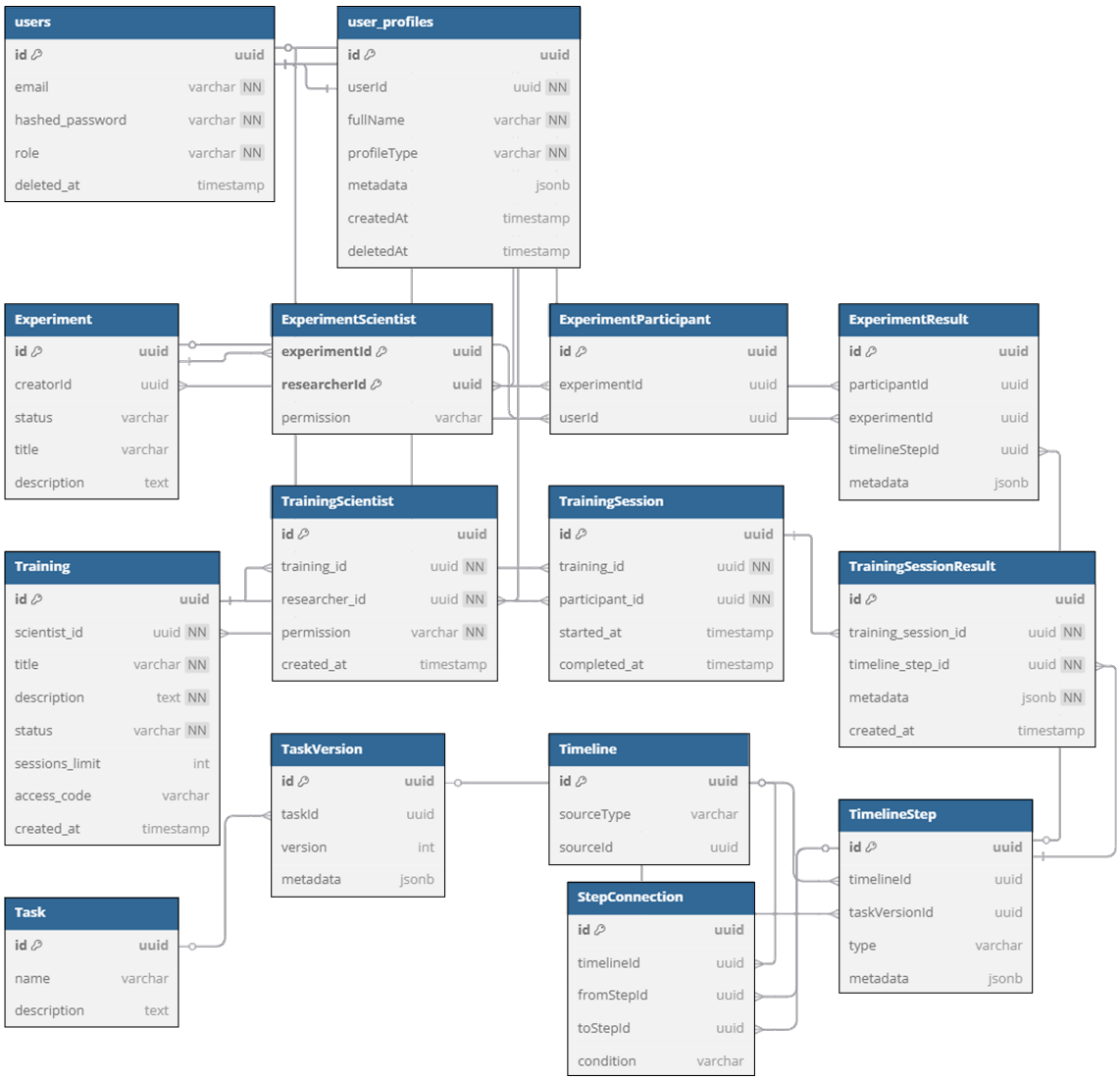

Figure 4.11: Complete conceptual data model of the CAPE-KIDS platform. The diagram illustrates the main entities and their relationships across the platform's core modules: user management, experiment and training workflows, task configuration, timeline structure, and result collection. Key elements include the separation of users and profiles, reusable task templates with versioning, flexible timelines with conditional logic, and support for both experimental and training sessions with result tracking.

Chapter 5

Implementation

This chapter details the technical implementation of the system specified in Chapter 4. It describes the concrete technology stack chosen for the project, the implementation of the backend and frontend components, and the specific software design patterns used to realize the specified functionalities. While Chapter 4 described *what* the system does, this chapter explains *how* it was built.

5.1 High-Level Component Architecture

The system is implemented using a classic client-server model, with three distinct tiers, as illustrated in Figure 5.1.

- **Frontend (Client):** A Single-Page Application (SPA) built with **React** and **TypeScript**. This component is responsible for rendering the entire user interface and, critically, for executing the entire experiment engine logic within the user's browser.
- **Backend (Server):** A RESTful API developed with **NestJS** (Node.js). Its primary responsibilities are user authentication, data persistence (communication with the database), and serving the experiment definitions (the timelines) to the client.
- **Database:** A **PostgreSQL** relational database that acts as the single source of truth, storing all user data, platform content, experiment configurations, and participant results.

5.2 Backend Implementation

The backend of the CAPE-KIDS platform was implemented as a monolithic RESTful API using the **NestJS** framework (v10), which runs on Node.js. NestJS was chosen for its modular architecture, native support for TypeScript, and its use of decorators and established design patterns, which facilitate the creation of scalable and maintainable server-side applications. The entire backend is responsible for user authentication, data persistence, and serving experiment definitions to the client application.

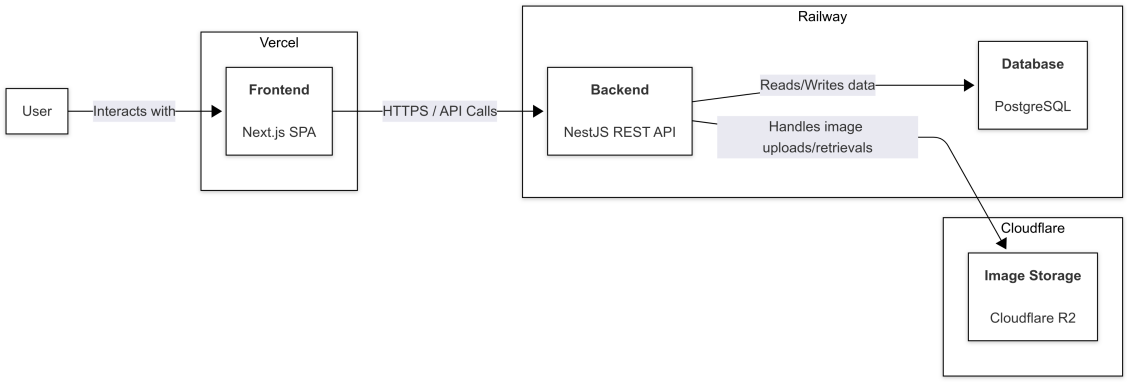

Figure 5.1: High-level component architecture of the CAPE-KIDS platform. The diagram illustrates how the main technological components are deployed and interact in production. The frontend (a Next.js SPA) is hosted on Vercel and communicates with a backend REST API (NestJS) deployed on Railway. The backend handles business logic, data persistence via a PostgreSQL database, and media storage operations via Cloudflare R2.

5.2.1 Data Persistence and ORM

The persistence layer was designed to map the conceptual data model specified in Chapter 4 to the **PostgreSQL** relational database. To manage all database interactions, the **TypeORM** library was selected as the Object-Relational Mapper (ORM). Each logical entity (e.g., ‘User‘, ‘Experiment‘, ‘Timeline‘) was implemented as a TypeORM entity class, using decorators to define columns, data types, and relationships (‘@OneToOne‘, ‘@OneToMany‘, ‘@ManyToOne‘). This approach allowed for a clear and maintainable mapping between the application’s object model and the database tables.

To implement the flexible metadata attributes required by the specification (e.g., in the ‘UserProfile‘ entity), the native `jsonb` data type from PostgreSQL was utilized. This choice provides the flexibility to store complex, nested JSON objects while also allowing for efficient querying of the data within those objects.

5.2.2 API Layer and Data Validation

The API endpoints were structured following REST principles. NestJS’s modular structure was leveraged to group related endpoints into modules (e.g., ‘UsersModule‘, ‘ExperimentsModule‘). For example, the ‘ExperimentsModule‘ contains controllers for handling HTTP requests related to experiments, services that encapsulate the business logic, and providers for interacting with the database via TypeORM.

A critical aspect of the backend implementation was ensuring robust data validation for all incoming requests and outgoing responses. To achieve this, the **Zod** validation library was adopted to define and validate the **Data Transfer Objects (DTOs)**. Instead of using traditional class-based DTOs with decorators, Zod schemas were used to define the shape and types of data for every API endpoint. This approach has several advantages:

- **Single Source of Truth:** The Zod schemas were placed in a shared, private library (‘capekids-shared‘), allowing both the NestJS backend and the React frontend to import the exact same validation rules. This guarantees end-to-end type safety and eliminates inconsistencies between the client and server.
- **Rich Validation:** Zod provides a rich and expressive API for defining complex validation rules that go beyond simple type checking.
- **Type Inference:** TypeScript types can be automatically inferred from the Zod schemas, reducing code duplication and the risk of type mismatches.

Figure 5.2 shows an example of the Zod schemas used to validate the different types of user profiles, which are then used to type-check the flexible ‘metadata‘ field.

```
export const researcherMetadataSchema = z.object({
  institution: z.string(),
  department: z.string(),
  degree: z.string(),
  researchAreas: z.array(z.string()).optional(),
  publications: z
    .array(
      z.object({
        title: z.string(),
        journal: z.string(),
        year: z.number(),
        link: z.string().optional(),
      })
    )
    .optional(),
  affiliations: z.array(z.string()).optional(),
  orcidId: z.string().url().optional(),
  linkedin: z.string().url().optional(),
  googleScholar: z.string().url().optional(),
});
```

```
export const psychologistMetadataSchema = z.object({
  licenseNumber: z.string().optional(),
  specialties: z.array(z.string()).optional(),
  certifications: z.array(z.string()).optional(),
  yearsExperience: z.number().optional(),
  clinicOrInstitution: z.string().optional(),
  assessmentTools: z.array(z.string()).optional(),
});
```

```
export const participantMetadataSchema = z.object({
  age: z.number(),
  gender: z.string(),
  handedness: z.string().optional(),
  nativeLanguage: z.string().optional(),
  medicalConditions: z.array(z.string()).optional(),
  medications: z.array(z.string()).optional(),
  parentalConsent: z.boolean(),
});
```

Figure 5.2: Example of Zod schemas implemented in the shared library to ensure type-safe validation of the `metadata` field for different user profiles.

To further enhance the development workflow, **Swagger UI** was integrated to provide interactive API documentation. Through a custom integration, the Zod schemas were used to automatically generate the OpenAPI (Swagger) specification. As a result, the interactive documentation not only lists all available endpoints but also provides accurate models and examples for request bodies derived directly from the Zod schemas, as shown in Figure 5.3. This ensures the API documentation is always synchronized with the validation logic.

5.2.3 Security and Testing

Security was a key consideration in the backend implementation. User authentication and session management are handled via stateless **JSON Web Tokens (JWT)**. Upon successful login, the backend signs and issues a JWT containing the user’s ID and role. This token must be sent by the client in the ‘Authorization‘ header of all subsequent requests to protected endpoints. A custom



Figure 5.3: Swagger UI documentation for an API endpoint, where the example schema is automatically generated from the corresponding Zod schema.

NestJS Guard was implemented to validate this token on every protected route, ensuring that users can only access the resources for which they have permission.

For automated testing, the **Jest** framework was used. A test suite was established with a focus on unit and integration tests for the backend services. While the goal is to achieve comprehensive coverage, due to time constraints, the testing efforts were prioritized on the most critical modules, such as the authentication and user management services, where full test coverage was implemented.

5.3 Frontend Implementation

The core logic of the platform, including the experimental engine and all user-facing interactions, resides in the frontend application. Implemented as a Single-Page Application (SPA) using **React** and **TypeScript**, it was designed to be modular, responsive, and maintainable.

5.3.1 State Management

To manage the application's complex state, the **Zustand** library was chosen for its simplicity and performance. A modular approach was taken, with separate stores dedicated to managing distinct parts of the application's state. This included dedicated stores for the timeline editing process, the runtime state of an active experiment (e.g., current step, score), and the state of specific UI components like modals and sidebars. This separation helped to keep the state logic organized and easy to debug.

5.3.2 Implementation of the Timeline Editor

The visual timeline editor, a central component of the platform, was implemented using the **React Flow** library. This library provides the foundation for building node-based interfaces.

The core of the editor is based on a dynamic node-edge data structure. Each step in a timeline is represented as a "node," and the connections between them are "edges." User interactions, such as reordering steps via drag-and-drop, trigger functions that update this underlying graph structure.

After the API request to save the new step order, the component re-renders to reflect the new experimental flow.

5.3.3 Implementation of the Step Configuration and Execution

Each step in the timeline is rendered within a canvas that maintains a fixed 16:9 aspect ratio, ensuring a consistent visual presentation regardless of screen size. The "Screen Editor" for custom blocks was implemented to allow researchers to dynamically add, position, and configure elements such as text and images via a drag-and-drop interface.

The various types of triggers (keyboard, mouse, timer) were implemented using the **Command design pattern**. Each trigger is an encapsulated object that continuously listens for its specific event during a trial. When the event occurs, it executes its defined action, such as advancing the timeline or stop the timeline execution. This decoupled approach makes the system highly extensible to new types of interactions.

5.3.4 The Experimental Engine: Implementation of the Timeline Model

The heart of the platform's engine is its **timeline-based** model. In this model, each experiment is structured as a timeline: a sequence of steps that defines the entire experimental flow from start to finish. This mirrors the way experimental protocols are typically designed in cognitive and developmental psychology.

The term "timeline-based" was deliberately chosen over the more generic "workflow-based" for several key reasons. It is more intuitive for researchers, it inherently suggests a process with a defined start and end, and it accurately reflects the specialized scope of executing self-contained research protocols.

From a software engineering perspective, the platform's core architecture was designed with a significant concentration of logic on the frontend (client-side). This means that the entire experimental engine—encompassing the timeline interpretation, task execution, and the handling of real-time participant triggers—operates within the user's browser. While this client-side approach does not follow a strict object-oriented paradigm, several architectural concepts and design patterns were leveraged to structure the application, ensure its extensibility, and manage its complexity.

- **Workflow Engine:** The system acts as a specialized **Workflow Engine**. The frontend "interprets" a timeline definition fetched from the backend, executing it step-by-step.
- **State Machine:** The experiment flow is managed as a finite **State Machine**, where each timeline step is a state and user actions trigger transitions to the next state in a predictable manner.
- **Command and Observer Patterns:** To handle user interactions, the **Command pattern** is used to encapsulate triggers as decoupled objects. The **Observer pattern** is then used

to ensure the user interface components react and re-render automatically in response to changes in the experiment's central state.

5.3.5 User Interface Implementation

The user interface was implemented as a series of React components, using the **Material-UI** (MUI) component library for a consistent visual design.

The visual Timeline Editor, a core component of the researcher's interface, was implemented using the **React Flow** library. This library provides the tools for building the specified interactive, node-based editor.

For the participant's interface, two critical technical choices were made to ensure data validity. First, a pre-loading mechanism was implemented to fetch all media assets before an experiment starts, preventing network lag during trials. Second, to measure reaction times with high precision, the browser's native **`performance.now()`** API was used.

5.3.6 Extensibility and Future Work

A key goal of the platform's implementation was to ensure that the core architecture is extensible, allowing for the incremental addition of new functionalities without requiring significant structural changes. Several design decisions were made to support this extensibility across the core modules.

Trigger System Modularity The trigger system was implemented using a registry-based Command pattern. Each trigger is defined as a standalone object, decoupled from the visual interface and execution logic. All available triggers are registered in a centralized `TriggerActionsRegistry`, and executed through a generic dispatcher function. This design allows new types of triggers (e.g., voice input, eye-tracking) to be added by simply implementing a new trigger handler module and registering it—without modifying the existing codebase.

Support for New Timeline Step Types Timeline steps are represented as strongly typed entities, each with a specific `type` field and corresponding metadata. The frontend uses a factory-based rendering system to determine which component should be used for each step type. This abstraction layer makes it easy to introduce new types of steps—such as conditional blocks, randomized sequences, or video-based stimuli—by simply extending the existing registry and defining the rendering and configuration logic.

Task Template Flexibility The task model is based on reusable timeline templates. As a result, researchers can create increasingly complex tasks by composing smaller, modular steps. This compositional structure ensures that even novel experimental paradigms can be implemented by reusing the existing architecture. In future iterations, support for branching logic, adaptive difficulty, and dynamic feedback can be added at the step level, without requiring a redesign of the timeline model.

Planned Enhancements Some future capabilities already envisioned include:

- Integration of survey instruments into the platform, allowing researchers to collect self-reported data as part of the experimental protocol.
- Integration with EEG devices to enable the collection of neurophysiological data synchronized with experimental events.
- Integration with wearable devices (e.g., accelerometer, gyroscope) to enable motion-based interactions.
- Import and export of task templates using a standardized JSON schema.

The underlying architecture was intentionally designed to accommodate these extensions in a scalable and maintainable manner.

5.4 User Interface Overview

This section presents the main interface components of the CAPE-KIDS platform. The user interface was designed to provide a smooth experience for researchers while preserving the underlying complexity of experimental logic. The platform centers the user experience around timeline configuration and task design, which are the most critical and expressive features of the system. For consistency and a professional look and feel, the UI components were built using the Material-UI library.

5.4.1 Researcher Interface and Workflow

The researcher's journey on the platform begins at the login screen (Figure 5.4), which ensures secure, role-based access to the system's functionalities. On this initial page, the user can also select their preferred language between English and Portuguese. This setting can be changed later on the configuration screen.

The application validates the user's session using a JWT token; if an expired token is detected, the user is automatically disconnected and redirected to this initial screen.

Once authenticated, the researcher is directed to their main dashboard. In its current implementation, the dashboard provides a high-level overview of the researcher's activity by displaying summary statistics, such as the total number of experiments, tasks, and participants on the platform.

From the dashboard or the main navigation menu, the researcher can access the core management areas: the list screens for Experiments, Trainings, and Tasks. A consistent design pattern is used across all three, providing a familiar and predictable user experience. Figure 5.5 shows the main management screen for Experiments, which serves as the central hub for a researcher's studies. This interface presents a clear, tabular view of all projects, allowing a user to search for a specific study, create a new one, or use the contextual action menu to perform key operations such

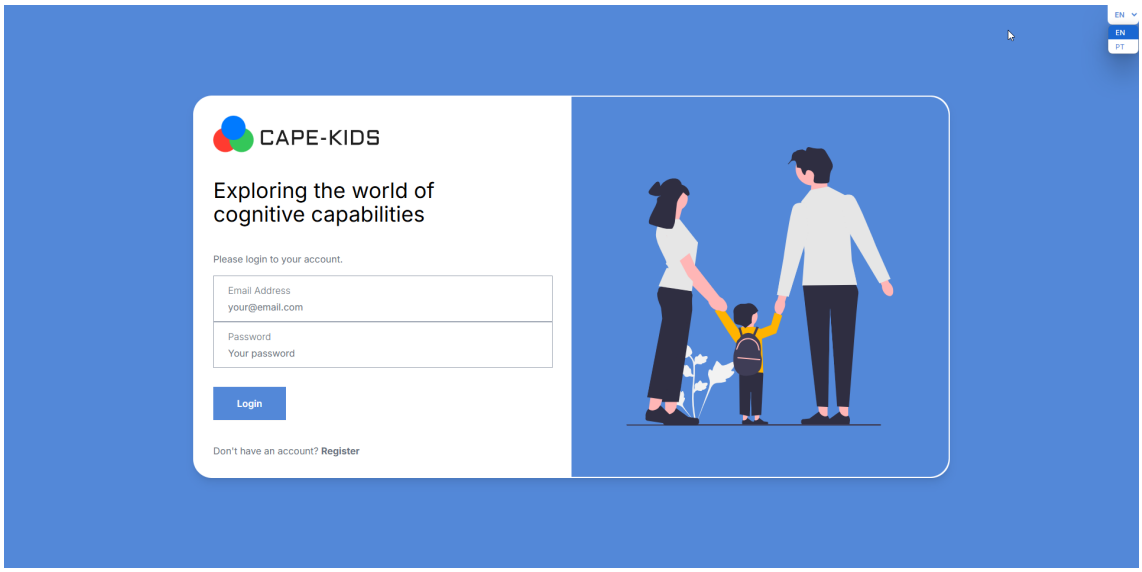

Figure 5.4: The platform's login screen, providing secure access and language selection.

as viewing the experiment's timeline or managing its participants. The corresponding interfaces for managing Trainings and reusable Tasks follow this exact same list-based design.

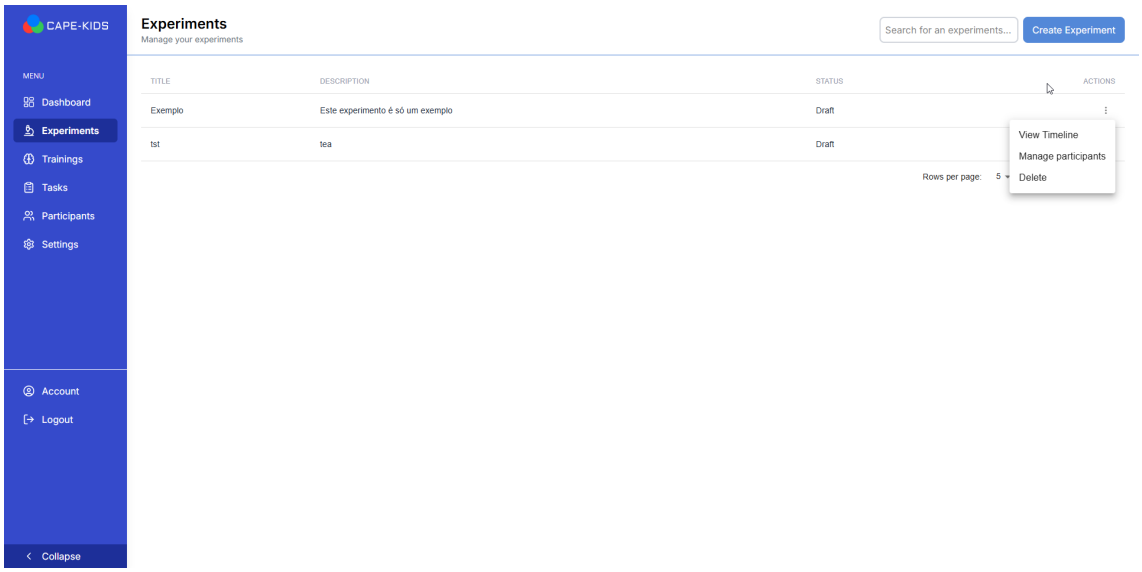

Figure 5.5: The main management interface for Experiments, demonstrating the list view and the contextual action menu for a study.

5.4.2 Visual Experiment Design: The Timeline Editor

After selecting a study from the management list, the researcher is directed to the Timeline Editor. This interface is the visual heart of the CAPE-KIDS platform, allowing for the creation and modification of the entire experimental flow. As shown in Figure 5.6, the editor displays the sequence

of events as a connected series of blocks, providing a clear, at-a-glance overview of the participant's journey. From here, the researcher can add new steps, reorder existing ones, and configure the properties of each block.

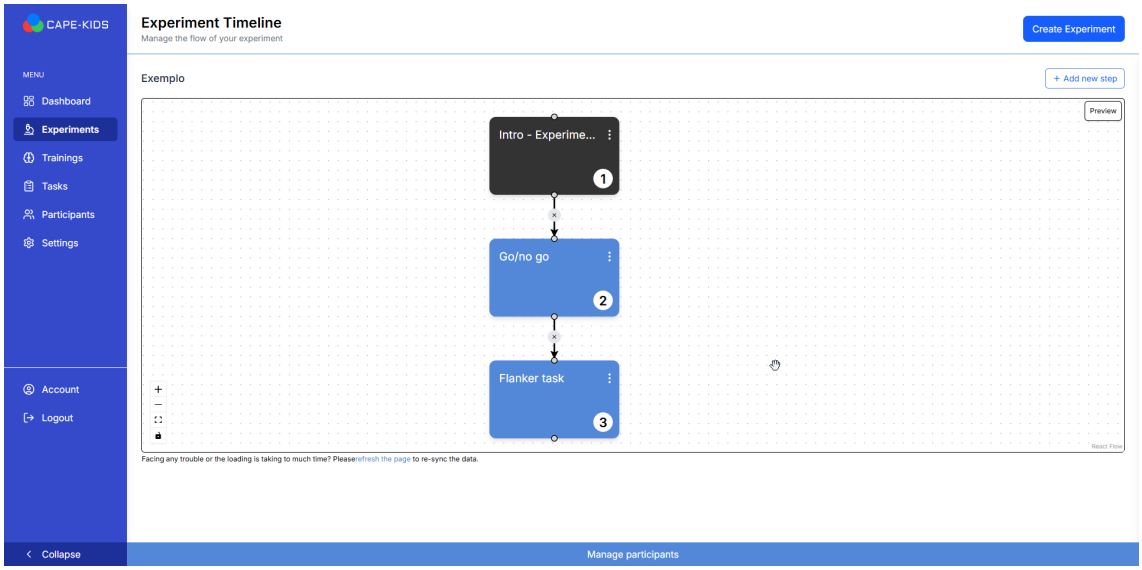

Figure 5.6: The main Timeline Editor, showing a simple, linear flow for an experiment.

A key architectural concept of the platform is that each "Task" block within a timeline is, itself, a self-contained, reusable timeline. This means a researcher can design a complex task once—complete with its own sequence of instructions, practice trials, and main trials—and save it as a master `Task` template. This template can then be dragged into any main experiment timeline as a single, modular block.

To illustrate this nested structure, Figure 5.7 shows the internal timeline of a "Go/No-Go" task template. It is clear that the task itself is a micro-flow with its own steps for instructions, a practice block, and the main task block. This entire sequence is what gets executed when this task is placed as a single item in a larger experiment's timeline, demonstrating the platform's composite and reusable design.

It is important to note that the Timeline Editor for the "Training" module follows this exact same design and functionality, providing a consistent and unified experience for the researcher regardless of whether they are creating an assessment or an intervention protocol.

5.4.3 Building Blocks: The Timeline Step Editor

The researcher builds an experiment's timeline by adding a sequence of "steps". The platform provides four distinct types of steps, each with a specific purpose: `Custom Block`, `Task`, `Sequential Stimuli`, and `Multi-Trigger Stimuli`.

The most fundamental building block is the **Custom Block**, typically used for creating instruction or feedback screens. As shown in Figure 5.8, this block type opens the "Screen Editor". This editor provides a canvas with a fixed 16:9 aspect ratio, ensuring a consistent visual layout

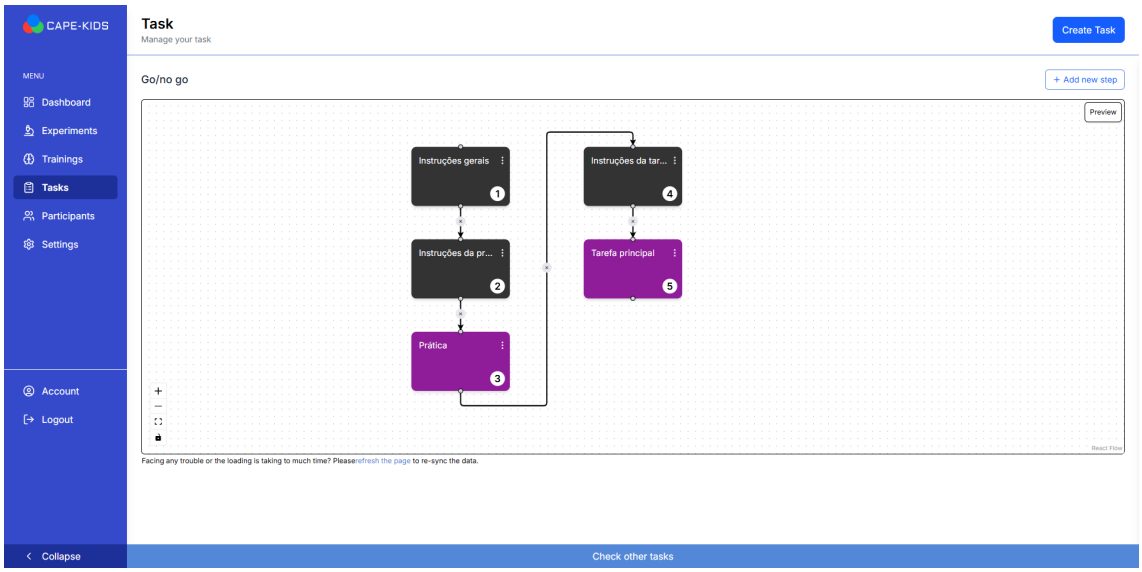

Figure 5.7: The expanded view of a "Go/No-Go" task template, showing its own internal, self-contained timeline.

across different screen sizes. Within this canvas, a researcher can add media elements like text and images, and define event **Triggers**—such as a keyboard press or a timer—which are linked to an action, like advancing to the next step in the timeline.

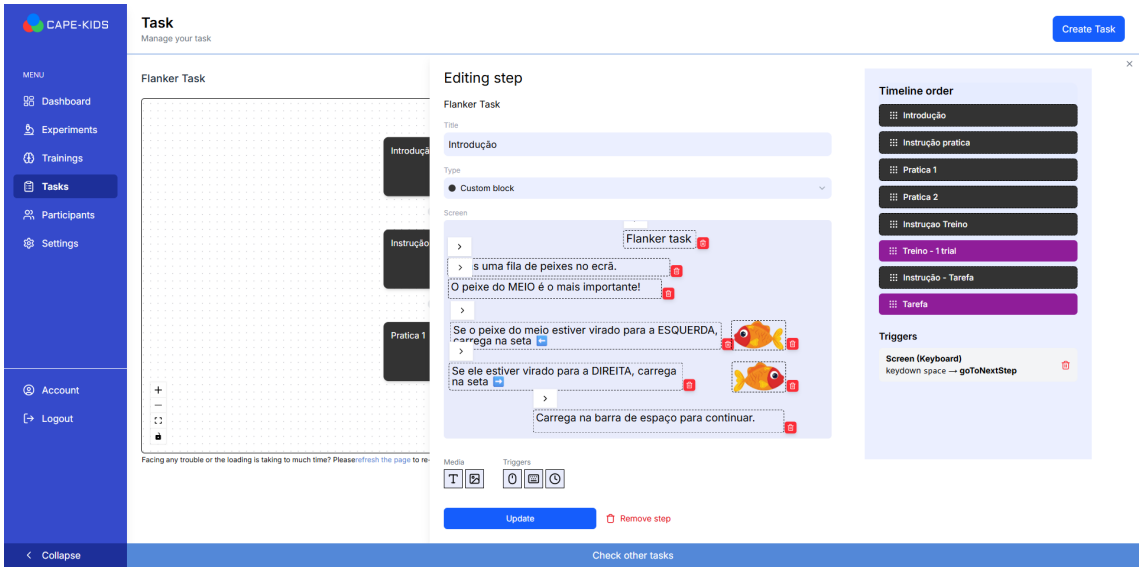

Figure 5.8: The configuration view for a `Custom Block`, showing the 16:9 screen editor with text, images, and a keyboard trigger defined.

The same core "Screen Editor" is used to build the stimuli for the more complex block types. A **Sequential Stimuli** block is a container designed for the main body of cognitive tasks where each stimulus requires an immediate, one-to-one response. This block has a global configuration modal (Figure 5.9) where parameters applicable to the entire block—such as the number of trials

or the default stimulus duration—can be defined.

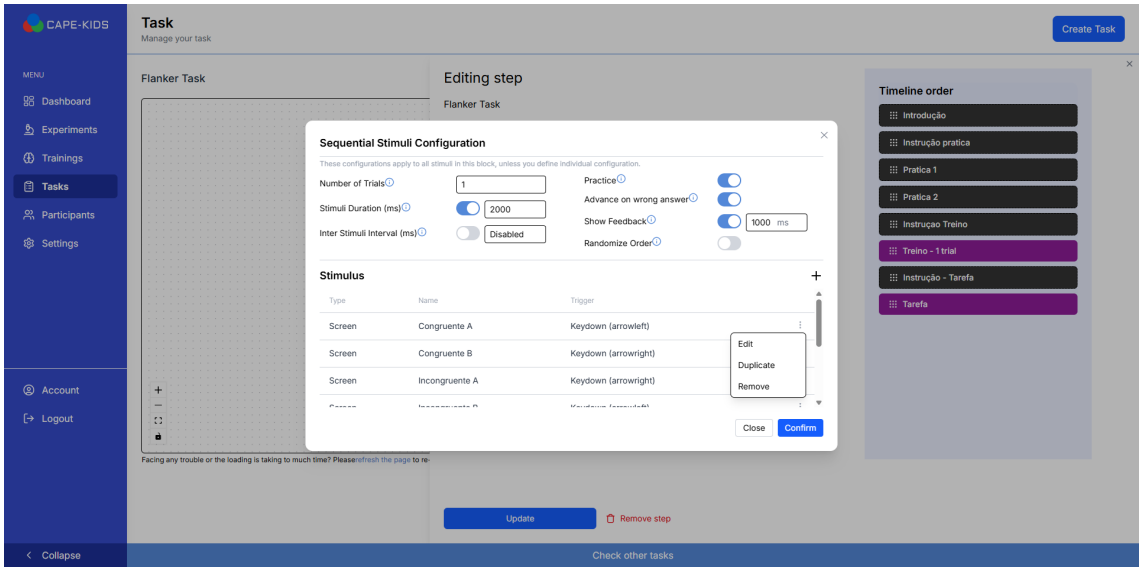

Figure 5.9: The configuration modal for a `Sequential Stimuli` block, showing block-level settings and a list of individual stimuli.

Within this block, each individual stimulus is configured on its own screen (Figure 5.10). While it inherits global settings like timing from the parent block, these can be individually overridden for specific trials. Critically, the response **Trigger** for each stimulus—such as the specific key press that defines a correct answer—is *always* defined at this individual level. This design ensures maximum flexibility, allowing each stimulus within a block to have a unique and specific interaction rule.

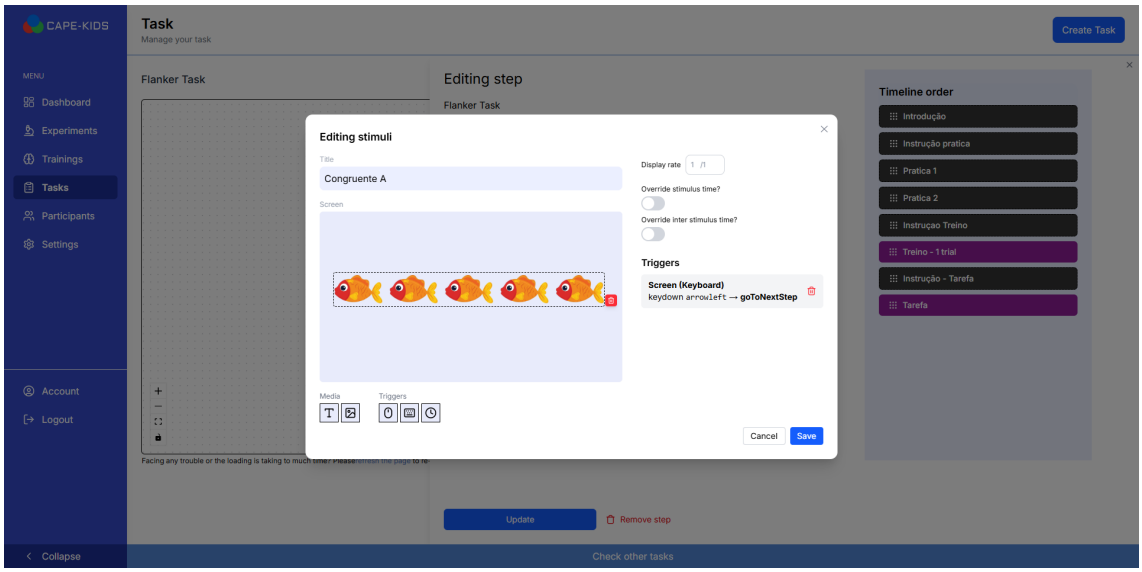

Figure 5.10: The editor for an individual stimulus, where its visual content and specific response trigger are configured.

To support tasks like working memory tests (e.g., Digit Span), the platform provides a **Multi-Trigger Stimuli** block. The key difference from a sequential block is the response logic: instead of a one-to-one response for each stimulus, the participant is shown a sequence of stimuli and must provide a single, combined response at the end.

Finally, the **Task** block serves a different purpose: it allows the researcher to insert a complete, pre-built task template (a reusable timeline) as a single, modular step, as described in the section on Task Versioning.

Once the experiment's timeline is designed, the next step in the workflow is to manage its participants. The platform provides a dedicated screen for each experiment (Figure 5.12) where the researcher can view all enrolled participants and add new ones by generating a unique access code for them to use.

5.4.4 Participant Management and Data Export

Once an experiment is designed, the researcher manages enrollment and data access through the "Experiment Participants" screen. This interface provides the unique `Access URL` and `Access Code` that can be shared with potential participants.

To enroll users in the study, the researcher can use the "Add participants" feature, which opens a modal window (Figure 5.11) displaying a list of available participants on the platform. After participants are added, they appear on the main list for that experiment. As shown in Figure 5.12, this screen allows the researcher to monitor enrollment and, crucially, to access the results. Once a participant has completed their session, a contextual action menu becomes available for that user, providing options to download their complete trial-by-trial data in standard formats, such as XLSX or JSON, ready for statistical analysis.

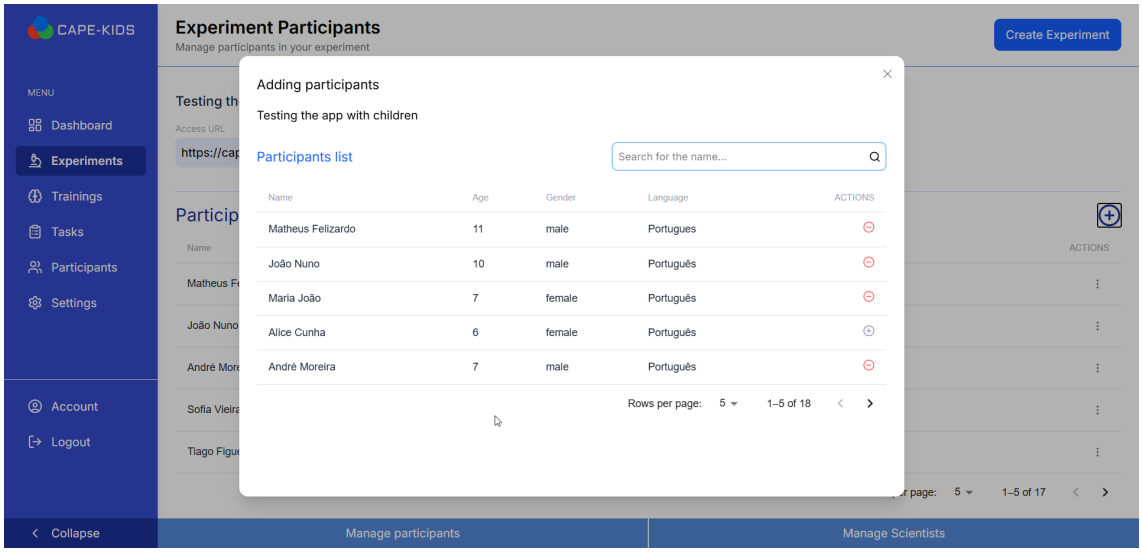

Figure 5.11: The interface for adding participants to a study from a list of available users on the platform.

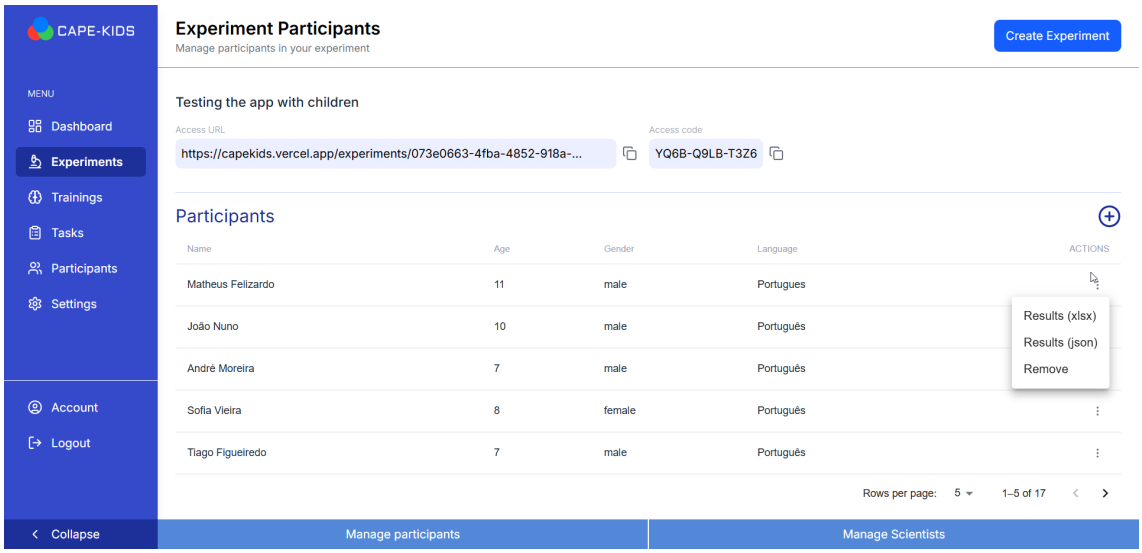

Figure 5.12: The participant management dashboard for a specific experiment. The open action menu illustrates the options to download the collected data for a completed session.

5.4.5 Participant Experience and Workflow

The interface and workflow for participants were designed to be simple, direct, and secure, minimizing distractions and focusing solely on the execution of the assigned protocols.

Accessing an Experiment A participant can access a study in two ways. First, after logging in, they are presented with a simplified dashboard that shows a list of only the experiments and trainings they are enrolled in. As shown in Figure 5.13, the interface is minimal, providing a clear action to "Start experiment".

Alternatively, a researcher can share a direct access URL for a specific study. The system ensures security by requiring the participant to be logged in and formally enrolled in the experiment to proceed. If either of these conditions is not met, the user is redirected to the login screen.

Experiment Execution Before the experiment begins, the platform pre-loads all necessary assets (such as images and sounds) for the entire timeline. This technical step is crucial to ensure a smooth, lag-free experience, preventing delays during critical task trials which could otherwise compromise the collected data.

Once loaded, the experimental engine begins executing the timeline, presenting each step sequentially to the participant. This includes instruction screens (Figure 5.14), task trials, and feedback screens, as designed by the researcher.

Completion and Data Integrity Upon completion of the final step in the timeline, the participant's results are automatically compiled and saved to the database. To maintain the scientific integrity of single-session experiments, the system prevents a participant from re-taking a study

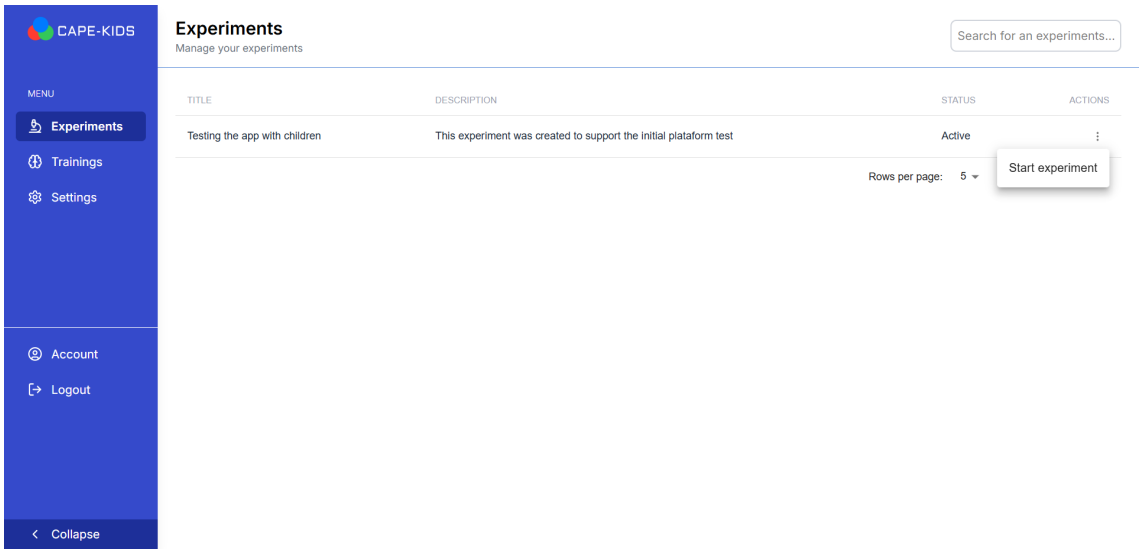

Figure 5.13: The participant's simplified dashboard, showing a list of enrolled experiments and the action to start a session.

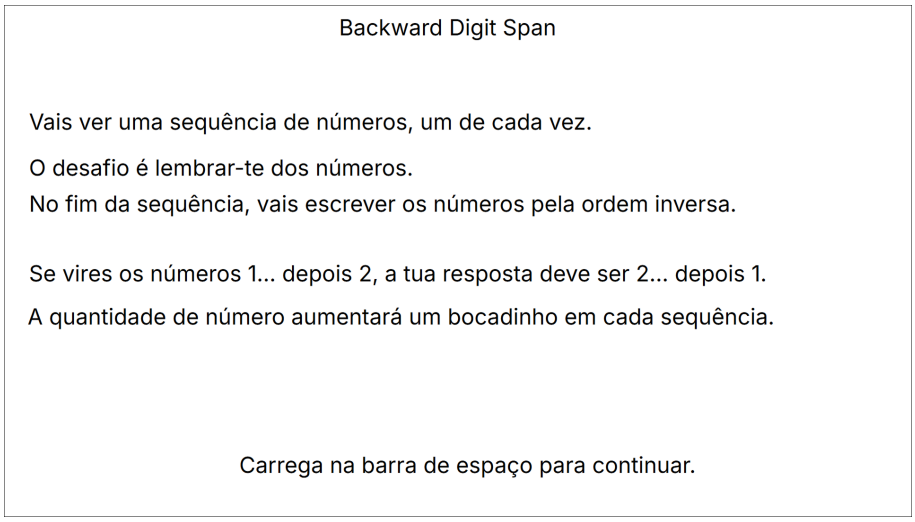

Figure 5.14: An example of an instruction screen presented to the participant before a task begins.

they have already completed. If they attempt to access the experiment again, they are presented with a notification message, as shown in Figure 5.15.

5.5 Chapter Summary

This chapter detailed the technical implementation of the CAPE-KIDS platform, explaining *how* the system was built based on the specification from Chapter 4. The architecture follows a classic client-server model with three tiers: a **React and TypeScript** frontend for the user interface and experimental engine, a **NestJS (Node.js)** backend for RESTful API services and data persistence, and a **PostgreSQL** database as the central data store.

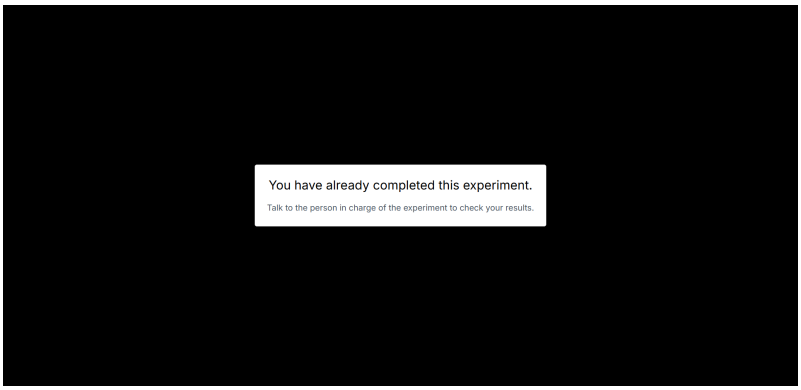

Figure 5.15: The notification message displayed to a participant attempting to re-access a completed experiment.

The backend implementation leveraged **NestJS's modular architecture**, employing **TypeORM** for Object-Relational Mapping to PostgreSQL, including the use of 'jsonb' for flexible metadata attributes. Robust data validation was achieved using **Zod schemas** in a shared library, ensuring end-to-end type safety and automatically generating interactive API documentation via **Swagger UI**. Security was managed through stateless **JSON Web Tokens**, and automated testing relied on the **Jest** framework for critical modules.

The frontend, a Single-Page Application, hosts the core experimental logic. It utilizes **Zustand** for modular state management and **React Flow** for the visual timeline editor, which handles a dynamic node-edge structure. Task steps are configurable within a fixed 16:9 canvas ratio, with event triggers implemented using the Command design pattern for extensibility. The experimental engine operates as a timeline-based Workflow Engine and State Machine, using Command and Observer patterns to manage participant interactions. For data validity, the frontend incorporates asset pre-loading and uses the browser's native 'performance.now()' API for high-precision reaction time measurements.

A key design principle was extensibility, achieved through modular trigger and timeline step systems, and flexible task templates. Envisioned future enhancements include integration with survey instruments, EEG devices, wearable technology, and standardized JSON schema import/-export.

Finally, the chapter provided an overview of the User Interface, detailing the researcher's workflow for experiment design, participant management, and data export, as well as the streamlined participant experience focused on secure and seamless experiment execution.

Chapter 6

Experimental Study: Procedure and Results

Following the design and implementation of the CAPE-KIDS prototype detailed in Chapter 4 and 5, the final phase of the research involved its validation. This chapter details the execution of a pilot experimental study conducted with the platform's target end-users: children. It outlines the procedure of this study—the participants, the materials used, and the data collection process—and presents the quantitative and qualitative results. Finally, it provides a discussion of these findings in relation to the platform's objectives and the core research question.

6.1 Usability and Feasibility Study with Children

A study with children was conducted to serve as a pilot test and initial validation of the CAPE-KIDS platform. The primary objective was not to generate new scientific findings on executive functions themselves, but rather to validate the platform's end-to-end usability and feasibility as a research tool: from the researcher's ability to create a complex experiment to the participant's ability to complete it, and critically, to verify the integrity and precision of the collected data.

6.1.1 Participants and Setup

Contact information was gathered for 17 potential participants. To streamline the process, user accounts were pre-created on the platform for each potential participant (e.g., `user1@capekids.com`) with a randomized password.

Due to time constraints, scheduling difficulties, and the logistical requirement of having a psychology specialist present for each session, the study was ultimately conducted with a final sample of **4 children** in the age range of the study (5 - 9 years old), Portuguese speakers. The participants, anonymized in this report, are referred to as P1, P2, P3 and P4. *Participation was voluntary and conducted only after obtaining explicit informed consent from the parents or legal guardians of each child, in accordance with ethical research guidelines.*

6.1.2 Materials: The Test Experiment

The material used was a single experiment created on the CAPE-KIDS platform. This experiment was designed with feedback from a child psychology specialist and used reduced and adapted versions of common EF tasks, suitable for testing the platform's functionalities. The experiment's timeline consisted of five main blocks presented in a fixed order:

1. Introduction Screen
2. Backward Digit Span Task
3. Go/No-Go Task
4. Flanker Task

The specific configuration of each cognitive task within the timeline was as follows:

Backward Digit Span The task began with instructions, followed by a 2-trial practice block. The main task consisted of 8 trials where the sequence length increased every 2 trials (from 2 numbers up to 4 numbers). An error would automatically advance the participant to the next task in the experiment's timeline.

Go/No-Go Task The task included an introduction, a practice block with 9 randomized trials (6 "Go", 3 "No-Go") with a 1000ms stimulus duration and feedback, and an instruction screen. The main task consisted of 24 trials with a 2:1 ratio of "Go" to "No-Go" stimuli. The stimulus was presented for 1000ms, followed by a 300ms Inter-Stimulus Interval (ISI).

Flanker Task This task was structured with an introduction, a 6-trial practice block (2 congruent, 2 incongruent, 2 neutral), and instructions before the main task. The main task contained 30 trials (10 of each condition: congruent, incongruent, neutral) presented in a randomized order, with a stimulus duration of 1500ms.

6.1.3 Procedure and Data Collection

Each session lasted approximately 15 minutes and was conducted in a quiet environment by a researcher specialist on executive functions. Observational notes were recorded manually during the session using a structured template, focusing on engagement, usability issues, and unexpected behaviors. The participants (P1, P2, P3, and P4) participated in the experiment once. The child interacted with the experiment on a desktop computer using a keyboard and mouse. The interaction between the child and the platform, along with the researcher's monitoring during the session, is depicted in Figure 6.1. It was noted by the specialist that the keyboard layout, particularly when a dedicated numeric keypad is absent and numbers are physically distant (e.g., on the top row), could potentially influence response times due to increased motor difficulty rather than cognitive processing time.

The CAPE-KIDS platform automatically recorded trial-by-trial quantitative data, including the stimulus presented, the participant's response, response time in milliseconds, and accuracy. For reaction time measurement, the platform's frontend leverages the browser's `performance.now()` API. This modern interface provides high-resolution, monotonic timestamps that are not subject to system clock adjustments, making it the current standard for performance measurement in web applications. The researcher also took qualitative observational notes on user engagement and any usability issues encountered.

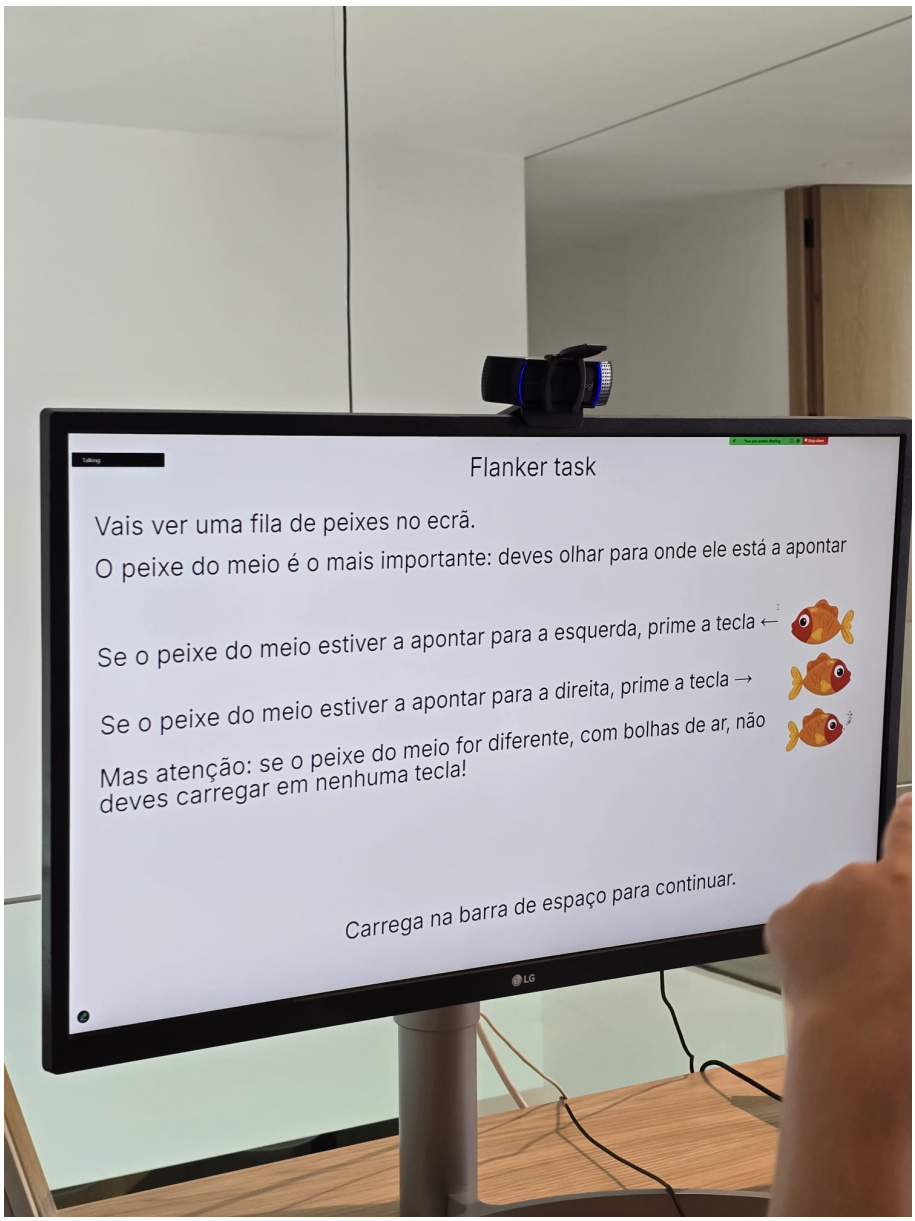

Figure 6.1: Instruction screen presented to children during the Flanker task (Portuguese version). Session monitored by a researcher specialist.

6.2 Results

The usability and feasibility study conducted with the child participants yielded positive initial results, confirming that the platform successfully meets its core objectives regarding data collection and user interaction.

6.2.1 Platform Feasibility and Data Integrity

The primary goal of validating the end-to-end workflow was achieved successfully. The platform was able to administer the entire multi-task timeline to all participants without technical failures. Upon completion of each session, the system correctly generated and saved the trial-by-trial data, which was successfully exported by the researcher in both XLSX and JSON formats for external analysis. This confirms the platform's feasibility as a data collection tool.

6.2.2 Qualitative Usability Findings

Observational notes taken during the sessions, complemented by post-session interviews with the children, indicated a high level of engagement and generally positive user experience.

- **Task Comprehension and Instructions:** Most children (P2, P3, P4) reported that the instructions for all tasks were easy to understand and they completed them successfully without needing help for comprehension. P1, however, found the Digit Span task easy to understand but experienced some difficulty with the Flanker Task instructions, requiring some help, and perceived the Go/No-Go and Flanker tasks to be fast.
- **Need for Assistance:** P3 and P4 completed all tasks without needing help. P2 needed help with the Flanker Task, while P1 required assistance across the Digit Span, Go/No-Go, and Flanker tasks. This highlights individual differences in task mastery and the need for adaptable support.
- **Engagement and Enjoyment:** All children expressed enjoyment in playing the games, confirming they were "fun". P1 described the games as "fun and annoying," suggesting a mix of challenge and engagement. P2 found it specifically "fun and entertaining". Favorite parts varied: P4 enjoyed Digit Span and Go/No-Go, P3 preferred Go/No-Go, while P1 and P2 favored the Digit Span task. All participants indicated a desire to play the games again.
- **Perceived Difficulty and Confusion:** The Flanker Task was consistently identified as the most difficult part by P2, P3, and P4. P1 also found the Go/No-Go and Flanker tasks difficult, specifically noting their speed. Most children (P1, P2, P3) reported feeling no confusion during the tasks, with P2 specifically finding the overall experience "funny". P1, however, expressed confusion regarding the Flanker Task.

- **Self-Perception of Performance:** P2, P3, and P4 all felt they were performing well, noting that they could identify when they made errors. P1, however, reported that she did not feel she was doing well, which aligns with her expressed difficulties.

6.2.3 Quantitative Performance Data from Child Participants

This section presents the individual performance data collected from the four child participants for each of the cognitive tasks.

6.2.3.1 P1

- **Backward Digit Span:** Advanced in training, correctly answering both 2-digit trials. In the main task, responded correctly to the first 2-digit trial but erred on the second, leading to progression to the Go/No-Go task. Average response time: 5540.75 ms (including training).
- **Go/No-Go:** In training (9 trials), erred 6 times and correctly answered 3. In the main task (24 trials), made 2 errors and 22 correct responses. Average response time for 14/16 correct 'Go' interactions: 613.64 ms.
- **Flanker:** In training (6 trials), erred 3 times and correctly answered 3. In the main task (30 trials), made 4 errors and 26 correct responses. Average response time for 17 correct actions: 975.11 ms.

6.2.3.2 P2

- **Backward Digit Span:** Erred on the second training trial (2 digits), leading to progression to the Go/No-Go task. Average response time: 5362.5 ms (including training).
- **Go/No-Go:** In training (9 trials), erred 3 times and correctly answered 6. In the main task (24 trials), made 2 errors and 22 correct responses. Average response time for 15/16 correct 'Go' interactions: 550.4 ms.
- **Flanker:** In training (6 trials), erred 3 times and correctly answered 3. In the main task (30 trials), made 2 errors and 28 correct responses. Average response time for 18 correct actions: 885.83 ms. One correct response was disregarded due to a reaction time of 90ms, likely indicating anticipation.

6.2.3.3 P3

- **Backward Digit Span:** Correctly answered all trials, including both 5-digit sequences, and advanced to the Go/No-Go task. Average response time: 6481.5 ms (including training). Notably, response time increased significantly for 5-digit sequences, reaching 14455 ms.

- **Go/No-Go:** In training (9 trials), made 0 errors and correctly answered all 9. In the main task (24 trials), made 0 errors and 24 correct responses. Average response time for 16/16 correct 'Go' interactions: 415.56 ms.
- **Flanker:** In training (6 trials), erred 3 times and correctly answered 3. In the main task (30 trials), made 3 errors and 27 correct responses. Average response time for 20 correct actions: 697.15 ms.

6.2.3.4 P4

- **Backward Digit Span:** Correctly answered all trials up to the first 5-digit attempt, where erred, leading to progression to the Go/No-Go task. Average response time: 7410.77 ms (including training). Notably, response time increased significantly for the 5-digit sequence, reaching 32463 ms, where the participant erred by transposing the last two digits.
- **Go/No-Go:** In training (9 trials), made 0 errors and correctly answered all 9. In the main task (24 trials), made 0 errors and 24 correct responses. Average response time for 16/16 correct 'Go' interactions: 486.56 ms.
- **Flanker:** In training (6 trials), erred 3 times and correctly answered 3. In the main task (30 trials), made 2 errors and 28 correct responses. Average response time for 20 correct actions: 722.25 ms.

Table 6.1: Summary of Quantitative Performance Data from Child Participants

Participant	Task	Outcome (Accuracy/Progression)	Avg. RT (ms)
P1	Backward Digit Span	Erred on 2-digits (main task)	5540.75
	Go/No-Go	22/24 correct	613.64
	Flanker	26/30 correct	975.11
P2	Backward Digit Span	Erred on 2-digits (training)	5362.5
	Go/No-Go	22/24 correct	550.4
	Flanker	28/30 correct	885.83
P3	Backward Digit Span	All correct (up to 5-digits)	6481.5
	Go/No-Go	All correct	415.56
	Flanker	27/30 correct	697.15
P4	Backward Digit Span	Erred on 5-digits (first attempt)	7410.77
	Go/No-Go	All correct	486.56
	Flanker	28/30 correct	722.25

6.3 Discussion

The pilot study with children, alongside the successful end-to-end workflow validation, provides initial evidence for the CAPE-KIDS platform's capability to effectively integrate assessment, training, and monitoring of executive functions in children aged 5 to 9.

6.3.1 Validation of Platform Objectives

The findings from this pilot study confirm the fundamental objectives of the CAPE-KIDS platform. The successful administration of a multi-task timeline to all participants without technical failures, coupled with the precise and complete collection of trial-by-trial data (including reaction times), validates the platform's feasibility as a robust and reliable tool for experimental data collection in a child population. The consistent data export functionality further underscores its utility for researchers seeking to perform detailed analysis.

Qualitatively, the high level of engagement and absence of significant usability issues observed indicate that the platform provides a positive and intuitive user experience for children within the target age group. This ease of use is crucial for maintaining focus during cognitive tasks and collecting valid data, aligning with the platform's design principles that prioritize accessibility and minimal distractions. The children's feedback, particularly their general enjoyment and desire to play again, reinforces the engaging nature of the game-like interface, which is vital for sustained participation in research and training protocols.

6.3.2 Addressing the Research Question: How can a web platform effectively integrate assessment, training, and monitoring of executive functions in children aged 5 to 9?

Based on the results of this pilot study, the CAPE-KIDS platform demonstrates a promising capacity to address the proposed research question:

- **Assessment:** The platform's ability to precisely record performance metrics (accuracy, errors, response times) for tasks assessing Working Memory (Backward Digit Span), Inhibitory Control (Go/No-Go), and Cognitive Flexibility/Attentional Control (Flanker) confirms its efficacy as an assessment tool. The varied performance levels observed across participants (e.g., P3 and P4's strong performance in Backward Digit Span and Go/No-Go vs. P1 and P2's more moderate performance) indicate the platform's sensitivity in capturing individual differences in EF. The detailed data, such as P4's specific error in digit sequence or P2's anticipatory response, provides rich, granular information crucial for diagnostic assessment. The qualitative data further supports this, as participants generally understood task rules and perceived their performance accurately, allowing for meaningful assessment.

- **Training:** While this pilot focused on assessment, the inherent modularity and configurability of the tasks, as highlighted in Chapter 4 (Specification) and Chapter 5 (Implementation), lay the groundwork for effective training protocols. The adaptive nature of the Backward Digit Span task (increasing difficulty with correct answers) exemplifies how the platform can be configured to provide individualized training, dynamically adjusting to a child's performance. The reusability of "Task" blocks as templates further supports the development of complex, multi-session training regimens. The children's positive engagement and desire

to play again suggest that the platform has the potential to sustain interest across multiple training sessions.

- **Monitoring:** The systematic collection and export of trial-by-trial data in standard formats (XLSX, JSON) are fundamental for long-term monitoring of executive functions. Researchers can leverage this data to track progress over multiple sessions, identify learning curves, and detect sustained changes in performance, which is vital for evaluating the effectiveness of training interventions or developmental trajectories. The platform's capacity to prevent re-taking completed sessions ensures data integrity for single-session experiments, a feature crucial for consistent monitoring. The observational notes regarding children needing help or finding tasks difficult provide valuable qualitative data that can complement quantitative metrics in a holistic monitoring approach.

6.3.3 Implications and Limitations

The initial results imply that CAPE-KIDS represents a viable and user-friendly web-based solution for researching executive functions in children. Its capacity to collect high-resolution data in a standardized manner, coupled with its flexible design, positions it as a valuable asset for the cognitive development research community.

However, it is crucial to acknowledge the limitations of this pilot study. The small sample size (four children) precludes drawing broad generalizations about typical performance or the statistical significance of observed differences. The study's primary objective was platform validation, not a comprehensive investigation into children's EF. Furthermore, this study involved only a single session per child, limiting insights into the platform's utility for long-term training and monitoring efficacy. The observation regarding keyboard layout potentially influencing response times highlights an external factor that more controlled laboratory settings or alternative input methods (e.g., touch screen) might mitigate in future, larger-scale studies aiming for extreme temporal precision.

6.3.4 Future Work

Building on the successful validation of the CAPE-KIDS platform, future research efforts should focus on several key areas. Larger-scale, longitudinal studies with a more diverse sample of children are necessary to fully validate the platform's training and monitoring capabilities over extended periods. This would involve evaluating the effectiveness of tailored training protocols and tracking developmental progress of executive functions. Further enhancements to the platform itself, as detailed in Chapter 5 (Extensibility and Future Work), including integration with survey instruments, EEG devices, and wearable technology, will expand its utility and enable more comprehensive research into the multifaceted nature of executive functions in children.

6.4 Chapter Summary

This chapter presented the pilot experimental study conducted with children to validate the CAPE-KIDS platform. The study's primary objective was to assess the platform's end-to-end usability and feasibility as a research tool, including data integrity and precision, rather than generating new scientific findings on executive functions.

The methodology involved a final sample of four Portuguese-speaking children (aged 5-9) participating in a single 15-minute session, remotely monitored by a psychology specialist. Participants interacted with a complex experiment on the platform, comprising adapted versions of Backward Digit Span, Go/No-Go, and Flanker tasks. Data collection included automatic recording of precise quantitative trial-by-trial data (response, reaction time using 'performance.now()') and qualitative observational notes and post-session interviews. It was noted that keyboard layout could potentially influence response times.

The results confirmed the platform's feasibility for data collection, demonstrating successful administration of the multi-task timeline and accurate data export. Qualitative findings indicated high engagement and generally positive user experience, with children reporting enjoyment, task comprehension (though some found Flanker challenging or fast, and some needed assistance), and varying self-perceptions of performance. Quantitative data provided individual performance metrics for each child across tasks, including accuracy, progression, and average reaction times, highlighting individual differences and specific task behaviors (e.g., significant RT increases for higher digit spans).

The discussion affirmed CAPE-KIDS's promising capacity to effectively integrate assessment (via precise data capture and sensitivity to individual differences), training (through modular, configurable, and adaptive task design), and monitoring (via systematic data collection for longitudinal tracking). While the study confirmed the platform's viability as a user-friendly web-based solution, limitations included the small sample size and single-session design. Future work should focus on larger, longitudinal studies, comprehensive automated testing, and integrating advanced features like physical activity-based tasks and audio instructions for younger children.

Chapter 7

Conclusion and Future Work

This dissertation addressed the persistent challenges in the assessment of executive functions (EFs) in children, namely the lack of standardization, accessibility, and innovation in existing tools. Through a phased research methodology, this work has successfully designed, implemented, and validated a novel solution: the CAPE-KIDS platform. This final chapter summarizes the main contributions of this project, acknowledges its limitations, and proposes directions for future work.

7.1 Conclusion and Main Contributions

This dissertation sought to answer the question of how the limitations of existing tools for EF assessment could be addressed through a new platform. The answer is provided through the design, implementation, and initial validation of CAPE-KIDS. The problem of **standardization** and **reproducibility** is addressed through the platform's immutable task versioning system. The challenge of **accessibility** for non-technical researchers is addressed by its visual, no-code timeline editor. Furthermore, the platform's **extensible architecture** provides a clear pathway to address the lack of innovation in the field, allowing for future integration of novel task types. The viability of this solution was confirmed through a feasibility study that demonstrated a successful end-to-end workflow, from experiment creation to successful data collection with child participants.

The primary contribution of this work is the **CAPE-KIDS platform**, a functional, web-based prototype designed as a direct response to the gaps identified in a systematic review of the scientific literature. The platform's main contributions can be summarized as follows:

- **A Standardized and Flexible Framework:** Through its timeline-based architecture and task versioning system, the platform provides a framework that promotes methodological consistency and scientific reproducibility, directly addressing the critical issue of non-standardization in the field.
- **An Accessible, No-Code Tool for Researchers:** The visual, intuitive interface empowers researchers without a technical background to create and manage complex cognitive experiments, lowering the barrier to entry for conducting digital EF research.

- **Successful Initial Validation:** The platform's design and functionality were validated through two distinct studies. A formative evaluation with subject-matter experts confirmed its scientific relevance and usability from a researcher's perspective. Subsequently, a usability and feasibility study with children demonstrated that the platform provides an engaging experience for its target end-users and can reliably collect performance data.

Furthermore, the research conducted for this dissertation resulted in significant scientific contributions, disseminated through peer-reviewed publications. The scoping review that formed the basis of this work was published [17], and more notably, the initial paper detailing the platform's architecture and design, titled "A Research-Oriented Platform for Standardized Executive Function Assessment and Training in Children" [7], was accepted for presentation at the **2025 IEEE 13th International Conference on Serious Games and Applications for Health (SEGAH)**. The work is scheduled to be published in the conference proceedings by IEEE, providing a strong peer-reviewed validation of the solution proposed in this dissertation.

7.2 Limitations

While this project successfully achieved its objectives, it is important to acknowledge its limitations to provide a balanced perspective on the work accomplished.

First, the usability study with children was conducted with a **small sample size** (N=4). Although sufficient for a pilot feasibility study, a larger and more diverse sample would be required to draw more generalizable conclusions about the platform's usability across different populations.

Second, the implemented prototype represents a **proof of concept**. While the core functionalities are robust, many advanced features envisioned—such as a comprehensive library of task templates or integrated data visualization tools—were outside the scope of this master's dissertation. In line with this, the project's focus on rapid prototyping meant that **automated test coverage**, while established, is not yet comprehensive. A testing strategy using the Jest framework was implemented for the backend, but complete coverage was only achieved for critical modules like authentication. This constitutes a technical debt that should be addressed to ensure the platform's long-term stability.

Finally, while the platform uses the high-precision `performance.now()` API for reaction time measurement, all browser-based timing is susceptible to external factors. The absolute millisecond-level accuracy was not benchmarked against laboratory-grade hardware, which remains a consideration for studies requiring the highest degree of temporal precision.

7.3 Future Work

The CAPE-KIDS platform was designed to be extensible, and this project serves as a solid foundation for several exciting directions for future work.

From a technical standpoint, a crucial next step is to implement a **comprehensive automated testing suite**. This would involve expanding the backend unit and integration test coverage using

Jest, and introducing an end-to-end (E2E) testing strategy for the frontend with a framework such as Cypress or Playwright. A robust testing foundation is essential before adding significant new features or deploying the platform in larger-scale studies.

Following this, the most immediate research step is to conduct a **larger-scale validation study** with children to gather more robust quantitative and qualitative data.

A primary direction for feature development is to fulfill the initial vision of **integrating physical activity-based tasks**. This would involve extending the platform to use device cameras, motion sensors and wearable devices. Importantly, the current architecture was designed from the ground up with this extensibility in mind, adopting a modular approach that allows new types of triggers and inputs to be integrated with minimal disruption to the core system. This foundation makes it feasible to implement hybrid tasks that combine cognitive demands with physical movements — such as jumping, gesture-based responses, or body orientation — opening possibilities for more engaging assessment protocols.

In parallel with these developments, another important feature is the integration of external neurophysiological hardware, such as **EEG (Electroencephalogram)** systems. This would enable the collection of brain activity data during task execution, addressing key suggestions from the expert validation phase and paving the way for more comprehensive cognitive monitoring and analysis.

Additionally, several refinements have been identified based on expert feedback and preliminary testing. These include the use of **audio-based instructions** to support pre-literate children (ages 5–7), the implementation of **interactive task demonstrations** before training blocks to ensure proper understanding, and the possibility of supporting **alternative response devices** (e.g., physical buttons or verbal responses) better suited to young children's motor and cognitive abilities. These improvements aim to enhance accessibility, engagement, and ecological validity in future studies.

Finally, other long-term enhancements include building a richer **library of task templates**, developing an **integrated dashboard for data analysis and visualization**, and exploring the platform's potential for **longitudinal training studies**, moving beyond single-session assessments to multi-session cognitive interventions.

In the long term, CAPE-KIDS can evolve into a comprehensive platform for digital cognitive research and training, supporting hybrid experimental designs that combine behavioral tasks, self-report instruments, and physiological data. With its modular architecture and no-code interface, it has the potential to democratize access to high-quality EF assessment tools across both research and educational settings.

References

[1] Sammy F. Ahmed, Jennie Grammer, and Frederick Morrison. Cognition in context: Validating group-based executive function assessments in young children. *Journal of Experimental Child Psychology*, 208:105131, 2021.

[2] Elien Bellon, Wim Fias, and Bert De Smedt. More than number sense: The additional role of executive functions and metacognition in arithmetic. *Journal of Experimental Child Psychology*, 182:38–60, 2019.

[3] John R. Best and Patricia H. Miller. A developmental perspective on executive function. *Child Development*, 81(6):1641–1660, 2010.

[4] Shi Yu Chan, Zi Yan Ong, Zhen Ming Ngoh, Yap Seng Chong, Juan H. Zhou, Marielle V. Fortier, Lourdes M. Daniel, Anqi Qiu, Michael J. Meaney, and Ai Peng Tan. Structure-function coupling within the reward network in preschool children predicts executive functioning in later childhood. *Developmental Cognitive Neuroscience*, 55:101107, 2022.

[5] Adele Diamond. Executive functions. *Annual Review of Psychology*, 64(Volume 64, 2013):135–168, 2013.

[6] Jill Fahy. Assessment of executive functions in school-aged children: Challenges and solutions for the slp. *Perspectives on School-Based Issues*, 15:151, 12 2014.

[7] Matheus Rodrigues Felizardo, António Coelho, Nuno Miguel Feixa Rodrigues, Eva Oliveira, Sónia Silva Sousa, and Adriana Sampaio. A research-oriented platform for standardized executive function assessment and training in children. In *Proceedings of the 5th International Workshop on Serious Games for Health and Medical Applications (SEGAM 2025)*, 2025. Paper submitted for publication.

[8] Ebru Ger and Florian J. Buehler. Is monitoring in executive functions related to metacognitive monitoring? *Cognitive Development*, 72:101514, 2024.

[9] Ebru Ger and Claudia M. Roebers. Monitoring and control processes within executive functions: Is post-error slowing related to pre-error speeding in children? *Journal of Experimental Child Psychology*, 246:105975, 2024.

[10] Amanda Grenell, Jacob R. Butts, Susan C. Levine, and Emily R. Fyfe. Children's confidence on mathematical equivalence and fraction problems. *Journal of Experimental Child Psychology*, 246:106003, 2024.

[11] Moeko Ishikawa, Yusuke Moriguchi, and Yasuhiro Kanakogi. Relationship between executive function and persistence in 5-year-olds. *Cognitive Development*, 67:101361, 2023.

[12] Dilara Keşşafoğlu, Aylin Küntay, and Berna A. Uzundağ. Immediate and delayed effects of fantastical content on children's executive functions and mental transformation. *Journal of Experimental Child Psychology*, 248:106067, 2024.

[13] Hedi Kwakkel, Mienke Droop, Ludo Verhoeven, and Eliane Segers. The impact of lexical skills and executive functioning on l1 and l2 phonological awareness in bilingual kindergarten. *Learning and Individual Differences*, 88:102009, 2021.

[14] Lydia Lavis and Caitlin E.V. Mahy. "i'll remember everything no matter what!": The role of metacognitive abilities in the development of young children's prospective memory. *Journal of Experimental Child Psychology*, 207:105117, 2021.

[15] K. Lertladaluck and Y. Moriguchi. Executive functions and theory of mind development in preschoolers: Insights from nirs data. *Neuropsychologia*, 205:109031, 2024.

[16] Brenna Lin, Jeffrey Liew, and Marisol Perez. Measurement of self-regulation in early childhood: Relations between laboratory and performance-based measures of effortful control and executive functioning. *Early Childhood Research Quarterly*, 47:1–8, 2019.

[17] Felizardo Matheus, Rodrigues Nuno, Coelho António, Sousa Sónia, Sampaio Adriana, and Oliveira Eva. Mapping executive function tasks for children: A scoping review for designing a research-oriented platform, 2025.

[18] Michelle N. Maurer and Claudia M. Roebers. Towards a better understanding of the association between motor skills and executive functions in 5- to 6-year-olds: The impact of motor task difficulty. *Human Movement Science*, 66:607–620, 2019.

[19] Akira Miyake and Naomi P. Friedman. The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 21(1):8–14, 2012. PMID: 22773897.

[20] Niamh Oeri and Claudia M. Roebers. Regulating disappointment can impair cognitive performance in kindergarten children: Individual differences in ego depletion. *Journal of Experimental Child Psychology*, 190:104728, 2020.

[21] Athanasia Papastergiou, Eirini Sanoudaki, Marco Tamburelli, and Vasiliki Chondrogianni. A study on the executive functioning skills of greek–english bilingual children – a nearest neighbour approach. *Bilingualism: Language and Cognition*, 26(1):78–94, 2023.

[22] Costanza Ruffini, Christian Tarchi, and Chiara Pecini. Which executive functions affect text comprehension and writing in paper and digital mode? an investigation in primary school children. *Computers & Education*, 207:104936, 09 2023.

[23] Katharina Schirmbeck, Nirmala Rao, Rhoda Wang, Ben Richards, Stephanie W. Y. Chan, and Claudia Maehler. Contrasting executive function development among primary school children from Hong Kong and Germany. *European Journal of Psychology of Education*, 36(4):923–943, December 2021.

[24] Jonas Schäfer, Timo Reuter, Miriam Leuchter, and Julia Karbach. Executive functions and problem-solving—the contribution of inhibition, working memory, and cognitive flexibility to science problem-solving performance in elementary school students. *Journal of Experimental Child Psychology*, 244:105962, 2024.

[25] Shelley Shaul, Mila Schwartz, Shelley Shaul, and · Schwartz. The role of the executive functions in school readiness among preschool-age children. *Reading and Writing*, 27, 08 2013.

[26] Phillip D. Tomporowski, Bryan McCullick, Daniel M. Pendleton, and Caterina Pesce. Exercise and children's cognition: The role of exercise characteristics and a place for metacognition. *Journal of Sport and Health Science*, 4(1):47–55, 2015.

[27] Laura Traverso, Irene Tonizzi, M. Carmen Usai, and Paola Viterbori. The relationship of working memory and inhibition with different number knowledge skills in preschool children. *Journal of Experimental Child Psychology*, 203:105014, 2021.

[28] Edy Veneziano and Eleonora Bartoli. Individual differences in the effectiveness of a narrative-promoting intervention: Relation with executive function skills. *First Language*, 42(4):523–547, 2022.

[29] Sha Xie, Chaohui Gong, Jiahao Lu, Hui Li, Dandan Wu, Xinli Chi, and Chunqi Chang. Enhancing chinese preschoolers' executive function via mindfulness training: An fnirs study. *Frontiers in Behavioral Neuroscience*, 16, 2022.

[30] Sha Xie, Shuqi Lu, Jiahao Lu, Chaohui Gong, and Chunqi Chang. Using mindfulness-based intervention to promote executive function in young children: a multivariable and multiscale sample entropy study. *Cerebral Cortex*, 34(9):bhae330, 09 2024.

[31] Theodore P. Zanto, Anastasia Giannakopoulou, Courtney L. Gallen, Avery E. Ostrand, Jessica W. Younger, Roger Anguera-Singla, Joaquin A. Anguera, and Adam Gazzaley. Digital rhythm training improves reading fluency in children. *Developmental Science*, 27(3):e13473, 2024.

[32] Jianhua Zhang, Jiacheng Lu, Youbin Sun, and Ji Li. Recreational gymnastics exercise of moderate intensity enhances executive function in chinese preschoolers: A randomized controlled trial. *PsyCh Journal*, 13(6):954–965, 2024.

Appendix A

Summary of Reviewed Executive Function Tasks

Table A.1: Summary of Reviewed Executive Function Tasks

Author(s)	Year	Task	Target EF	Objective	Stimulus	Trials
Ahmed et al. [1]	2020	Freeze Task	Inhibitory Control	Children march in a circle and freeze when music starts. They unfreeze when instructed or when music stops.	Music played for random intervals (7–15 seconds)	5 trials (Stop time: Reaction to freezing and distractibility)
		Jumping Task	Inhibitory Control, Working Memory	Children follow one- to three-step instructions after freezing (e.g., jump three times, clap twice).	Music played for random intervals (7–15 seconds)	3 trials (Stop time, distractibility, action recall, action performance)
		Marching Task	Inhibitory Control, Cognitive Flexibility	Children march in one direction for a specific song and the opposite direction for another song.	Music played for random intervals (7–15 seconds)	7 trials (Stop time, distractibility, march recall, march performance)
		Head-Toes-Knees-Shoulders	Inhibitory Control	Children perform the opposite of the instructed movement (e.g., touch toes when asked to touch head).	Not specified	30 trials across 3 blocks
		Backward Digit Span	Working Memory	Children repeat numbers in reverse order.	Spoken sequences of numbers	Not specified

Continued on next page

Author(s)	Year	Task	Target EF	Objective	Stimulus	Trials
		Pair Cancellation Task	Inhibitory Control	Children identify specific pairs (e.g., a ball followed by a dog) on a stimulus sheet.	Rows of pictures on a sheet	Not specified
Bellon et al. [2]	2019	Flanker Task	Inhibitory Control	Assess inhibition by responding to a central arrow's direction while ignoring flanking distractors.	Arrows flanked by distractors, presented for 2500 ms and black screen visible for 1000 ms	40 test trials (50% incongruent), preceded by 12 baseline trials with single arrows.
		Animal Stroop Task	Inhibitory Control	Assess inhibition by identifying the larger animal in real life, ignoring image size.	Two animals presented simultaneously, with one image four times larger. Stimuli appeared for 2000 ms, followed by a black screen that remained visible for 1000 ms	40 test trials (50% incongruent), preceded by 20 baseline trials with equal-sized images.
		Wisconsin Card Sorting Task (WCST)	Cognitive Flexibility	Measure cognitive flexibility by sorting cards based on changing rules (color or shape).	Cards with specific colors and shapes.	50 trials; sorting rule changes after 7–9 correct responses.

Continued on next page

Author(s)	Year	Task	Target EF	Objective	Stimulus	Trials
		2-Back Task	Working Memory	Measure updating by identifying if a current stimulus matches one shown two trials earlier.	Colored images, presented for 3000 ms each followed by black screen that remained visible for 1000 ms	40 test trials, with 30% being targets.
Chan et al. [4]	2022	Wisconsin Card Sorting Test (WCST)	Cognitive Flexibility	Measure cognitive flexibility by switching rules in categorization tasks.	Cards with attributes of color, shape, and number; rules change after 10 consecutive correct responses.	64 total trials.
Ger et al. [8]	2024	Hearts and Flowers Task	Inhibitory Control, Cognitive flexibility	Assess inhibition and cognitive flexibility through rule-switching.	Geometric shapes (hearts and flowers); heart = congruent, flower = incongruent.	Hearts: 24, Flowers: 36, Mixed: 48 (with 12 flowers).
Ger et al. [9]	2024	Hearts and Flowers Task	Inhibitory Control, Cognitive Flexibility	Assess inhibitory control and cognitive flexibility.	Stimuli presented on a screen: Hearts (congruent) and Flowers (incongruent). Each trial has 750 ms stimulus duration, followed by a 500 ms fixation cross.	Block 1: 24 congruent trials; Block 2: 36 incongruent trials; Block 3: 48 congruent + 12 incongruent trials.

Continued on next page

Author(s)	Year	Task	Target EF	Objective	Stimulus	Trials
Grenell et al. [10]	2024	Flanker Task	Inhibitory Control	Measure inhibitory control and attention.	Five fish or arrows; central target's direction is identified (e.g., «««< or «>«).	Not specified; includes congruent and incongruent trials.
		List Sorting Task	Working Memory	Measure working memory by recalling sequences of items by category and size.	Visual presentation of items (e.g., animals, foods).	6 levels (2–7 items).
		Dimensional Change Card Sort (DCCS)	Cognitive Flexibility	Measure cognitive flexibility by switching between sorting by shape or color.	Cards with objects of varying shapes and colors.	Practice trials followed by test trials (number not specified).
Ishikawa et al. [11]	2023	Red/Blue Task	Inhibitory Control	Measure inhibitory control through responses to incongruent color names.	Cards colored red or blue; children must point to the opposite color when named by the experimenter.	10 test trials (5 red, 5 blue).
		Bear/Dragon Task	Inhibitory Control	Measure inhibitory control by following commands from the "bear" and inhibiting those from the "dragon".	Commands issued by two puppets (bear and dragon).	12 trials (6 per puppet).

Continued on next page

Author(s)	Year	Task	Target EF	Objective	Stimulus	Trials
		DCCS Task	Cognitive Flexibility	Assess cognitive flexibility through rule-switching in card sorting.	Cards with two dimensions (color and shape); sorted based on alternating rules.	8 trials per phase (Pre-, Post-Switch, Mixed).
Keşşafoğlu et al. [12]	2024	Backward Word Span	Working Memory	Measure working memory by recalling words in reverse order.	Verbal sequences of 2–5 words (e.g., "curtain, garden").	8 test trials.
		Simon Says Task	Inhibitory Control	Assess inhibitory control via selective imitation and non-imitation of commands.	Actions (e.g., touch your nose) given with or without "Simon Says."	10 trials (5 imitation, 5 anti-imitation).
		Flexible Item Selection	Cognitive Flexibility	Measure cognitive flexibility by selecting cards matching on varying dimensions.	Three cards per trial; select two matching on one dimension, then another.	15 trials.
Kwakkel et al. [13]	2021	Head-Toes-Knees-Shoulders (HTKS)	Inhibitory Control, Working Memory	Measure self-regulation by responding oppositely to commands.	Verbal commands like "Touch your head"; responses must be opposite (e.g., touch toes).	20 test items (10 per rule set).
		Continuous Performance Task (CPT)	Inhibitory Control	Measure sustained attention using go/no-go responses.	Visual task: target = red circle, non-target = red square; auditory task: high/low tones.	200 trials per task (80 target, 120 non-target).

Continued on next page

Author(s)	Year	Task	Target EF	Objective	Stimulus	Trials
		Forward Word Span	Working Memory	Measure working memory by recalling word sequences.	Spoken words in sequences of 2–7.	2 trials per sequence length.
Lavis et al. [14]	2021	Simon Says Task	Inhibitory Control	Assess inhibition by following commands only when preceded by "Simon Says."	Verbal instructions to perform actions, sometimes preceded by "Simon Says."	10 trials (5 inhibition trials).
		Backward Word Span Task	Working Memory	Measure working memory by repeating verbal sequences in reverse order.	Words presented verbally by the experimenter.	2 trials per sequence length (2–5 words).
Lertladaluck et al. [15]	2024	Stroop Task	Inhibitory Control	Measure cognitive inhibition by responding to shape rather than color.	Shapes (circle, triangle, square) shown in congruent/incongruent colors; response via touchscreen.	6 blocks (18 trials each: 9 congruent, 9 incongruent).
		Luria Hand Game	Inhibitory Control	Measure inhibition through congruent/incongruent hand gestures.	Experimenter makes gestures (fist/point); child must mimic or do opposite depending on the session.	8 trials per session (congruent/incongruent).

Continued on next page

Author(s)	Year	Task	Target EF	Objective	Stimulus	Trials
		Flanker Task	Inhibitory Control	Assess inhibition and selective visual attention by ignoring adjacent distractors.	Fish symbols pointing congruently or incongruently to the central target.	20 trials (10 congruent, 10 incongruent).
Lin et al. [16]	2019	Shape Stroop Task	Cognitive Flexibility, Inhibitory Control	Measure cognitive flexibility and inhibition by identifying subdominant visual features.	Cards displaying shapes embedded within other shapes.	24 trials.
		Snack Delay Task	Inhibitory Control	Measure inhibitory control and delay of gratification using edible stimuli.	Transparent cup covering an M&M candy; delay durations of 10, 20, 30, and 15 seconds.	4 trials.
		Toy Delay Task	Inhibitory Control	Measure inhibitory control and delay of gratification using non-edible stimuli.	Transparent cup covering "silly bandz"; delay durations of 10, 20, 30, and 15 seconds.	4 trials.
Maurer et al. [18]	2019	Flanker Task	Inhibitory Control	Measure inhibition by focusing on a target while ignoring distractors.	Visual stimuli: Fish pointing in congruent or incongruent directions.	Block 1: 20 congruent; Block 2: 20 congruent + 20 incongruent.
		Animal Updating Task	Working Memory	Measure working memory updating by recalling the last two animals shown.	Sequences of animals presented for 1900 ms each, followed by a question mark.	2 test blocks of 5 trials each.

Continued on next page

Author(s)	Year	Task	Target EF	Objective	Stimulus	Trials
		Advanced Dimensional Change Card Sort (DCCS)	Cognitive Flexibility	Measure cognitive flexibility by sorting cards by color, shape, or a switching rule.	Cards with colors and shapes; instructions change across blocks (e.g., sort by shape if star is present).	Block 1: 10 color; Block 2: 10 shape; Block 3: 20 mixed (shape if star present).
Oeri et al. [20]	2020	Minnesota Executive Function Scale (MEFS)	Cognitive Flexibility, Working Memory, Inhibitory Control	Assess cognitive flexibility, working memory, and inhibition through card sorting tasks.	Cards sorted by changing rules (color vs. shape); increasing difficulty across 7 levels.	5 trials per part (A and B); adaptive difficulty.
		Flanker Task	Inhibitory Control	Measure response inhibition by identifying the orientation of a central target amidst distractors.	Fish stimuli: congruent (all fish face the same way) or incongruent (target faces opposite).	24 trials (12 congruent, 12 incongruent).
Papastergiou et al. [21]	2022	Nonverbal Stroop Task	Inhibitory Control	Measure inhibition by identifying arrow directions irrespective of position.	Arrows pointing in different directions; congruent and incongruent blocks.	3 blocks of 60 trials.
		Backward Digit Span Task	Working Memory	Assess updating by recalling numbers in reverse order.	Auditory presentation of number sequences.	4–8 numbers per trial, increasing gradually.

Continued on next page

Author(s)	Year	Task	Target EF	Objective	Stimulus	Trials
		Colour-Shape Task	Inhibitory Control, Cognitive Flexibility	Evaluate shifting between color and shape categorization tasks.	Shapes (circle/triangle) in different colors (red/blue) presented in top or bottom halves.	3 blocks: 32 trials (single-task blocks) + 64 trials (mixed-task block).
Schäfer et al. [24]	2024	Go/No-Go Task	Inhibitory Control	Assess inhibition by responding to go stimuli and withholding for no-go stimuli.	Animal images (cow = No-Go, chicken/pig/donkey = Go) presented sequentially for 500 ms each.	60 Go trials, 20 No-Go trials.
		Corsi Blocks Backward Task	Working Memory	Measure working memory by recalling sequences of highlighted blocks in reverse order.	3x3 grid; fairy images highlight blocks in increasing sequence lengths (1–8 blocks).	6 sequences per length.
		Flexible Item Selection Task (FIST)	Cognitive Flexibility	Assess cognitive flexibility through categorization shifts based on new shared properties.	Images sharing one or two properties (e.g., color, shape).	18 one-commonality trials; 9 two-commonalities trials.
Schirmbeck et al. [23]	2021	Child-Friendly Stroop Task	Inhibitory Control	Assess cognitive inhibition by overriding automatic responses.	Colored shapes, fruits, and vegetables; congruent and incongruent color pairings.	4 sets (20 items per set).

Continued on next page

Author(s)	Year	Task	Target EF	Objective	Stimulus	Trials
		Object Span Task	Working Memory	Measure working memory by recalling sequences of objects while performing a secondary task.	Cards with images of objects (e.g., apple, basket); "edible or not?" secondary task.	5 trials with spans increasing (2–6 objects).
Traverso et al. [27]	2021	Keep Track Task	Working Memory	Measure WM updating by recalling the last items in designated categories.	Pictures from five categories (animals, sky, fruit, vehicles, clothes).	6 trials: first 3 with one category; last 3 with two categories.
		Dual Request Selective Task	Working Memory	Assess visual-spatial WM by recalling positions while performing a concurrent task.	Frog pathway on a 4x4 chessboard; clap when frog jumps on a red square.	10 trials, increasing difficulty (pathway lengths = 2–6).
		Simon Says Task	Inhibitory Control	Measure inhibitory control by performing actions only when prompted by “Simon says.”	Verbal commands with gestures.	10 trials.
		Preschool Matching Familiar Figure Task	Inhibitory Control	Assess impulsive behavior by comparing target figures to alternatives.	Target figures with five alternatives on a page.	14 trials.
Veneziano et al. [28]	2022	Animal Stroop Task	Inhibitory Control	Assess inhibition by naming animal bodies while ignoring faces.	Boards with congruent, control-face, and incongruent animal drawings.	3 boards (24 items each).

Continued on next page

Author(s)	Year	Task	Target EF	Objective	Stimulus	Trials
		Classification Task	Cognitive Flexibility	Measure cognitive flexibility by categorizing objects by different criteria.	9 wooden pieces differing in shape, color, and size; 3 containers for sorting.	3 sorting trials.
		Local/Global Task	Cognitive Flexibility	Assess flexibility in recognizing global vs. local shapes.	Boards with geometric figures containing smaller shapes within their outlines.	6 boards.
Xie et al. [30]	2024	Dimensional Change Card Sort (DCCS)	Cognitive Flexibility	Assess cognitive flexibility by sorting cards based on changing rules.	Cards with two attributes (color and shape).	4 blocks (8 randomized trials per block).
		Go/No-Go Task	Inhibitory Control	Measure inhibition by responding to "Go" stimuli and withholding for "No-Go."	Images of animals (e.g., cows, horses, tigers); dog = "No-Go."	3 blocks (16 Go, 4 No-Go trials per block).
		Working Memory Task	Working Memory	Measure ability to recall visual-spatial locations.	A monkey appearing on one of six trees; recall the location after the image disappears.	3 blocks (8 trials each).
Xie et al. [29]	2022	Dimensional Change Card Sort (DCCS)	Cognitive Flexibility	Assess cognitive shifting by sorting cards based on changing rules.	Target cards (red boat, blue rabbit) vs. test cards (blue boat, red rabbit).	3 sessions, each with pre- and post-switch phases (25s each).

Continued on next page

Author(s)	Year	Task	Target EF	Objective	Stimulus	Trials
		Go/No-Go Task	Inhibitory Control	Measure inhibitory control by responding to "Go" stimuli and withholding for "No-Go."	Animal images (cow, horse, tiger = "Go"; dog = "No-Go").	3 sessions (10 Go trials, 10 No-Go trials per session).
		Missing Scan Task	Working Memory	Measure working memory by identifying missing objects after a brief presentation.	Sets of 4 animals displayed for 10s, disappearing into a "house"; 3 animals reappear.	3 sessions (5 trials each).
Zanto et al. [31]	2024	Flanker Task	Inhibitory Control	Assess inhibitory control by identifying the central arrow direction amidst distractors.	Five-arrow arrays (congruent/incongruent).	28 trials (50% congruent/incongruent).
		Stroop Task	Inhibitory Control	Measure inhibitory control by identifying text color instead of word meaning.	Words representing colors (e.g., "red") displayed in congruent or incongruent colors.	40 trials (50% congruent/incongruent).
		Working Memory Task	Working Memory	Measure working memory via reverse Corsi block-tapping task.	Squares illuminated in sequences, requiring participants to recall them in reverse order.	Adaptive trials (set size increases after correct trials).
Zhang et al. [32]	2024	Stop Signal Task	Inhibitory Control	Evaluate inhibition by stopping responses to a "beep" signal.	Left/right arrows with "beep" stop signal.	48 trials (16 per condition).

Continued on next page

Author(s)	Year	Task	Target EF	Objective	Stimulus	Trials
		Letter Memory Task	Working Memory	Assess working memory updating.	Letter sequences displayed, requiring sequential recall.	8 trials.
		Keep Track Task	Working Memory	Evaluate working memory by tracking recent items in categories.	Visual stimuli from 5 categories (e.g., colors, animals) displayed in octets.	9 trials.
		Dots Task	Cognitive Flexibility	Assess cognitive flexibility with rule switching.	Sun/moon stimuli requiring same-side or opposite-side responses.	20 trials.

Appendix B

Executive Function Tasks Categorized by EF Component

Table B.1: Executive Function Tasks Categorized by EF Component

Task	Author(s)	Year	Age Range/Mean	Notes
Inhibitory Control - 18 tasks				
Stroop Task **(7)**	Bellon et al. [2]	2019	6-8 years	Animal Stroop Task
	Lertladaluck et al. [15]	2024	5–7 years	
	Lin et al. [16]	2019	≈ 5 years	Shape Stroop Task
	Papastergiou et al. [21]	2022	5–9 years	Nonverbal Stroop Task
	Schirmbeck et al. [23]	2021	8–10 years	Child-Friendly Stroop Task
	Veneziano et al. [28]	2022	6—8 years	Animal Stroop Task
	Zanto et al. [31]	2024	8–9 years	
Flanker Task **(6)**	Bellon et al. [2]*	2019	6—8 years	
	Grenell et al. [10]	2024	7–9 years	
	Lertladaluck et al. [15]	2024	5–7 years	
	Maurer et al. [18]	2019	5–6 years	
	Oeri et al. [20]	2020	5–6 years	
	Zanto et al. [31]	2024	8–9 years	
Simon Says Task **(3)**	Keşşafoğlu et al. [12]	2024	5–6 years	

Continued on next page

Task	Author(s)	Year	Age Range/Mean	Notes
	Lavis et al. [14]	2021	≈ 5 years	
	Traverso et al. [27]	2021	5–6 years	
Go/No-Go Task **(3)**	Schäfer et al. [24]	2024	6—8 years	
	Xie et al. [30]	2024	5–6 years	
	Xie et al. [29]	2022	5–6 years	
Hearts and Flowers Task **(2)**	Ger et al. [8]	2024	6—8 years	This task also involves cognitive flexibility.
	Ger et al. [9]	2024	6—8 years	
Head-Toes-Knees-Shoulders **(2)**	Ahmed et al. [1]	2020	5–6 years	This task also involves working memory.
	Kwakkel et al. [13]	2021	5–6 years	
Freeze Task **(1)**	Ahmed et al. [1]	2020	5–6 years	
Jumping Task **(1)**	Ahmed et al. [1]	2020	5–6 years	
Pair Cancellation Task **(1)**	Ahmed et al. [1]	2020	5–6 years	
Red/Blue Task **(1)**	Ishikawa et al. [11]	2023	5–6 years	
Bear/Dragon **(1)**	Ishikawa et al. [11]	2023	5–6 years	
Continuous Performance Task **(1)**	Kwakkel et al. [13]	2021	5–6 years	
Luria Hand Game **(1)**	Lertladaluck et al. [15]	2024	5–7 years	
Snack Delay Task **(1)**	Lin et al. [16]	2019	≈ 5 years	
Toy Delay Task **(1)**	Lin et al. [16]	2019	≈ 5 years	
Preschool Matching Familiar Figure Task **(1)**	Traverso et al. [27]	2021	5–6 years	
Stop Signal Task **(1)**	Zhang et al. [32]	2024	5–6 years	
Marching Task **(1)**	Ahmed et al. [1]	2020	5–6 years	
Working Memory - 12 tasks				
Backward Digit/Word Span **(4)**	Ahmed et al. [1]	2020	5–6 years	
	Papastergiou et al. [21]	2022	5–9 years	
	Keşşafoğlu et al. [12]	2024	5–6 years	
	Lavis et al. [14]	2021	≈ 5 years	

Continued on next page

Task	Author(s)	Year	Age Range/Mean	Notes
Working Memory Task **(2)**	Xie et al. [30] Zanto et al. [31]	2024 2024	5–6 years 8–9 years	
Keep Track Task **(2)**	Traverso et al. [27] Zhang et al. [32]	2021 2024	5–6 years 5–6 years	
Forward Word Span **(1)**	Kwakkel et al. [13]	2021	5–6 years	
Animal Updating Task **(1)**	Maurer et al. [18]	2019	5–6 years	
2-Back Task **(1)**	Bellon et al. [2]	2019	6—8 years	
Object Span Task **(1)**	Schirmbeck et al. [23]	2021	8–10 years	
Corsi Blocks Backward Task **(1)**	Schäfer et al. [24]	2024	6—8 years	
Dual Request Selective Task **(1)**	Traverso et al. [27]	2021	5–6 years	
Missing Scan Task **(1)**	Xie et al. [29]	2022	5–6 years	
Letter Memory Task **(1)**	Zhang et al. [32]	2024	5–6 years	
List Sorting Task **(1)**	Grenell et al. [10]	2024	7–9 years	
Cognitive Flexibility - 8 tasks				
Dimensional Change Card Sort (DCCS) **(5)**	Grenell et al. [10] Ishikawa et al. [11] Maurer et al. [18] Xie et al. [30] Xie et al. [29]	2024 2023 2019 2024 2022	7–9 years 5–6 years 5–6 years 5–6 years 5–6 years	
Wisconsin Card Sorting Task (WCST) **(2)**	Bellon et al. [2] Chan et al. [4]	2019 2022	6—8 years ≈ 5 years	
Flexible Item Selection **(2)**	Keşşafoğlu et al. [12] Schäfer et al. [24]	2024 2024	5–6 years 6—8 years	
Colour-Shape Task **(1)**	Papastergiou et al. [21]	2022	5–9 years	
Classification Task **(1)**	Veneziano et al. [28]	2022	6—8 years	

Continued on next page

Task	**Author(s)**	**Year**	**Age Range/Mean**	**Notes**
Local/Global Task **(1)**	Veneziano et al. [28]	2022	6—8 years	
Dots Task **(1)**	Zhang et al. [32]	2024	5–6 years	
Minnesota Executive Function Scale (MEFS) **(1)**	Oeri et al. [20]	2020	5–6 years	This task also involves working memory and inhibitory control.