Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 292 (2026) 7763–7772

Procedia Computer Science

www.elsevier.com/locate/procedia

30th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2026)

A Multimodal Artificial Intelligence Framework for Early Detection of Impulse Control Disorders in Children and Adolescents

Daniel Vélez Gutiérrez[a,*], Giulia Maria Mattia[a], Ralf Seepold[a], Julian D. Echeverry-Correa[b], Jorge Enrique Avila Campos[b], Polina Mihova[c], Margarita Stankova[c], Kostadin Angelov[d], Natividad Martínez Madrid[e], Juan Antonio Ortega[f]

[a]*IAF Institute, HTWG Konstanz, Alfred-Wachtel-Str. 8, Konstanz, 78462, Germany*
[b]*Universidad Tecnológica de Pereira, Pereira, Colombia*
[c]*Department of Health Care and Social Work, New Bulgarian University, Sofia, Bulgaria*
[d]*Department of Surgery, Medical University Sofia, "Aleksandrovska", Bulgaria*
[e]*School of Informatics, Reutlingen University, Reutlingen, Germany*
[f]*University of Seville, Seville, Spain*

Abstract

Impulse control disorders (ICDs) in children and adolescents are difficult to identify at an early stage because symptoms are often subtle, heterogeneous, and context-dependent. Delayed recognition may contribute to long-term behavioral, psychological, and social consequences. Recent progress in artificial intelligence, mobile sensing, and wearable technologies offers new opportunities to collect behavioral and physiological signals in everyday environments and to derive digital behavioral biomarkers for early screening. This paper presents a framework and study design contribution; empirical validation is the subject of an ongoing pilot study. The framework *ImpulseAI* integrates language and communication data, sleep-related parameters, and stress-related physiological signals collected via mobile and wearable technologies to enable early, low-cost detection of impulse-related behavioral risks in children and adolescents. These heterogeneous data streams are processed through feature extraction, multimodal data fusion, and machine learning methods to identify patterns associated with impulsivity and behavioral dysregulation. A central contribution of the proposed approach is the combination of linguistic, sleep, and physiological indicators within an interpretable AI pipeline. Explainable artificial intelligence methods are considered to support transparency and to make model outputs understandable for clinical, psychological, and educational use. The framework is intended as a decision-support tool rather than as a replacement for clinical diagnosis. A pilot study design is outlined to evaluate the feasibility of data acquisition, feature extraction, and multimodal classification in a pediatric population. The expected outcome is a scalable and non-invasive screening approach that may support earlier assessment, preventive intervention, and long-term monitoring in pediatric mental health care.

© 2026 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of the KES International Association

Keywords: Multimodal Artificial Intelligence; Impulse Control Disorders; Digital Behavioral Biomarkers; Language and Sleep Analysis

* Daniel Vélez Gutiérrez. Tel.: +49-7531-206-9165
E-mail address: jvelezg@htwg-konstanz.de

1877-0509 © 2026 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of the KES International Association

1. Introduction

Impulse control disorders (ICDs) comprise a group of neuropsychiatric conditions characterized by difficulties in regulating impulsive behaviors and inhibiting inappropriate actions. In children and adolescents, increased impulsivity is frequently associated with behavioral dysregulation, emotional instability, and difficulties in social interaction. Elevated impulsivity has been linked to several developmental and psychiatric conditions, including attention deficit hyperactivity disorder (ADHD), conduct disorders, and substance-related disorders [1, 2]. Early manifestations of impulsive behavior may therefore represent important indicators of emerging mental health problems during childhood and adolescence. However, early detection of impulse-related disorders remains a significant challenge in clinical practice. Conventional diagnostic procedures rely primarily on clinical interviews, behavioral observations, and standardized questionnaires administered by clinicians, teachers, or caregivers. While these approaches provide valuable insights, they are often subjective and limited in their ability to capture subtle behavioral changes in everyday environments. In addition, diagnostic procedures typically require considerable time and specialized expertise, which may delay early identification and intervention [3].

Recent advances in digital health technologies and artificial intelligence (AI) have created new opportunities for continuous behavioral monitoring and early detection of neuropsychiatric conditions. Smartphones, wearable devices, and ambient sensing systems enable the collection of behavioral and physiological data in natural environments. Such data can be transformed into *digital biomarkers*, defined as objective, quantifiable measurements of human behavior or physiological processes obtained via digital devices [4]. In particular, non-invasive sensor technologies and wearable systems have increasingly been investigated for monitoring physiological signals related to sleep, respiration, and cardiovascular activity. These technologies enable continuous health monitoring outside clinical settings, thereby supporting the development of scalable digital health solutions [5]. Previous studies have demonstrated the feasibility of contactless, unobtrusive sensor systems for measuring cardiorespiratory parameters during sleep, enabling long-term monitoring of physiological signals in home environments.

Language and speech characteristics represent another promising source of behavioral information. Linguistic features such as lexical diversity, semantic coherence, and speech rhythm can reflect underlying cognitive and emotional processes. Advances in natural language processing and machine learning have therefore enabled automated analysis of speech and textual data as potential indicators of mental health conditions [6]. Sleep behavior is also closely linked to emotional regulation and impulse control. Insufficient or disrupted sleep has been associated with increased emotional reactivity and impulsive decision-making, particularly during adolescence [7]. Continuous monitoring of sleep patterns using wearable technologies may therefore provide valuable insights into behavioral regulation processes. Against this background, multimodal artificial intelligence approaches are gaining increasing attention in digital health research. By integrating heterogeneous data sources such as language signals, physiological measurements, and contextual behavioral information, multimodal AI models can identify complex patterns that may not be detectable through single modalities.

This paper introduces *ImpulseAI*, a multimodal artificial intelligence framework designed to support the early and low-cost detection of impulse control disorders in children and adolescents. The proposed approach integrates automated language analysis with sleep monitoring and stress-related physiological signals collected through mobile and wearable technologies. By identifying digital behavioral biomarkers associated with impulsive behavior, the framework aims to support early screening and preventive interventions in pediatric mental health care. The main contribution of this paper is the conceptual design of a multimodal AI framework that integrates language-based behavioral indicators with physiological and contextual data to facilitate early identification of impulse-related behavioral risks. To achieve this, the paper introduces a comprehensive framework that integrates data from multiple sources, including language, sleep, and physiological data. Additionally, it outlines a conceptual pipeline for extracting digital behavioral biomarkers, which is essential for understanding impulsivity. The paper also proposes a pilot study design to detect pediatric impulsivity, ensuring the research is geared toward practical application. Furthermore, the integration of explainable AI is emphasized to provide interpretable risk assessments, allowing stakeholders to better understand and act upon the identified risks.

2. Background and Related Work

Impulse control disorders (ICDs) and related behavioral dysregulation phenomena are closely associated with the psychological construct of impulsivity, commonly defined as a predisposition toward rapid and poorly considered reactions without adequate evaluation of potential negative consequences [2, 8]. In children and adolescents, elevated impulsivity is frequently observed in neuro-developmental and behavioral disorders such as attention-deficit hyperactivity disorder (ADHD) and conduct disorders. These conditions are associated with increased risk for long-term psychosocial difficulties, including academic problems, substance abuse, and emotional dysregulation. Contemporary psychological models further emphasize that impulsive behavior is strongly influenced by emotional states. In particular, the urgency framework explains how intense positive or negative emotions may trigger rash actions and risky behaviors, thereby linking affect regulation with impulsive decision-making processes [9]. Early identification of impulse-related behavioral patterns is therefore an important objective in pediatric mental health research and clinical practice.

Traditional diagnostic approaches primarily rely on clinical interviews, behavioral observations, and standardized questionnaires administered by clinicians, teachers, or caregivers. Although these methods provide valuable insights, they remain inherently subjective and often fail to capture subtle behavioral variations in everyday environments [3]. In addition, access to specialized clinical expertise may be limited, potentially delaying diagnosis and early intervention. These limitations have stimulated increasing interest in digital health technologies that enable continuous monitoring of behavioral indicators outside clinical settings.

Recent advances in mobile sensing, wearable technologies, and artificial intelligence have led to the emergence of *digital biomarkers*, defined as objective and quantifiable behavioral or physiological measurements collected through digital devices such as smartphones or wearable sensors [4]. Digital biomarkers have shown considerable potential for monitoring mental health conditions and supporting early detection of behavioral changes. In particular, multimodal digital biomarker approaches that combine physiological signals, behavioral indicators, and contextual data are increasingly explored to improve the robustness and interpretability of predictive models [10].

Language and speech analysis represent one promising modality for behavioral assessment. Variations in linguistic complexity, speech rhythm, lexical diversity, and semantic coherence can reflect underlying cognitive and emotional processes. Advances in natural language processing and machine learning have enabled automated detection of mental health indicators from speech and textual data, opening new opportunities for objective behavioral monitoring [6]. Jiang et al. [11] present a robust multimodal framework that integrates facial, vocal, linguistic, and cardiovascular features collected via remote interviews, and report validated classification results against clinical reference measures. While this work establishes a strong methodological foundation for analyzing multimodal mental health biomarkers, it focuses on adult populations in structured interview settings. Notably, it does not incorporate sleep as a modality, raising the question of whether continuous, naturalistic monitoring of sleep and physiological signals could enhance language-based indicators in everyday environments.

Additionally, most existing approaches focus on single modalities or adult populations, highlighting a gap in multimodal pediatric-focused solutions. Despite these technological advances, the application of multimodal artificial intelligence for the early detection of impulse control disorders in children remains limited. Existing studies often focus on a single data modality or on adult populations. A recent protocol study has examined the relationship among sleep, stress, and impulsivity in children and adolescents and proposed a scalable data-collection design that combines questionnaires, specialist evaluation, mobile sensors, and speech analysis. In their work, Scherz et al. [12] address impulse control disorders in youth directly and demonstrate the feasibility of a scalable, multimodal data-collection design combining mobile sensors, speech analysis, specialist evaluation, and standardized questionnaires in a pediatric population—the closest empirical precedent to the present work. That study, however, does not formalize a multimodal fusion architecture or incorporate an explainability layer, and its machine-learning analysis does not treat sleep monitoring as a dedicated modality. This work builds on Schertz et al. by integrating language analysis, sleep monitoring, and stress-related physiological signals into a multimodal AI approach for the early, low-cost identification of risk patterns associated with impulse control disorders. The framework is currently at the conceptual and design stage.

The following section introduces the conceptual architecture of the proposed *ImpulseAI* framework and describes how heterogeneous behavioral and physiological signals can be integrated for multimodal analysis.

3. ImpulseAI Framework

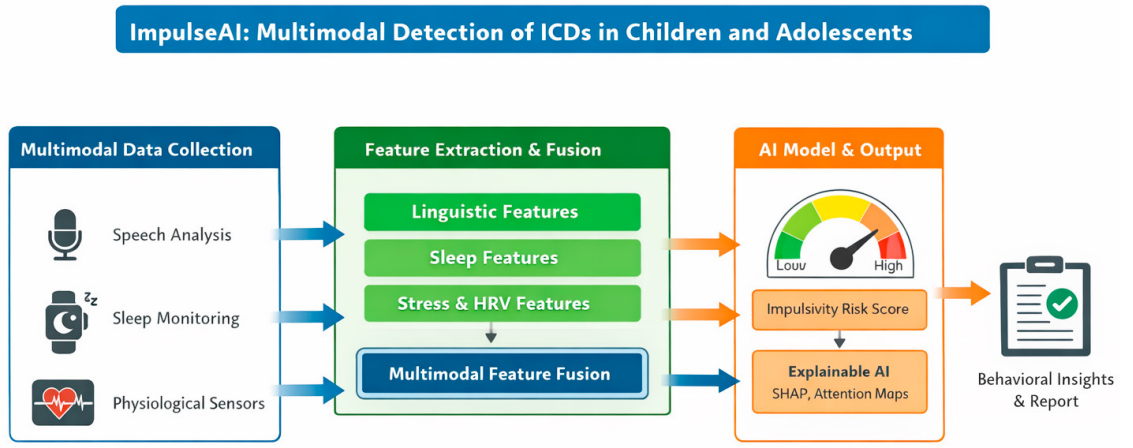


Fig. 1. Overview of the *ImpulseAI* framework integrating multimodal data sources (speech, sleep, and physiological signals) for early detection of impulse control disorders.

The *ImpulseAI* framework is conceived as a multimodal artificial intelligence architecture for the early detection of behavioral patterns associated with impulse control disorders in children and adolescents. The framework combines diverse data from digital devices and wearable technologies to identify behavioral indicators that may suggest the onset of impulse-related dysregulation. The components outlined below represent the design of ImpulseAI. This paper does not present experimental results; instead, it aims to validate these components in the pilot study detailed in Section 4. Conceptually, the framework comprises four main components, as shown in Fig. 1, and it integrates: (1) data acquisition, (2) feature extraction, (3) multimodal data fusion and machine learning, and (4) behavioral risk assessment. Together, these components form a processing pipeline that transforms behavioral and physiological signals into interpretable risk indicators. Also, Table 1 maps all ten pipeline components across these four framework stages.

3.1. Data Acquisition

The first component focuses on collecting behavioral and physiological signals using commonly available digital technologies. Three primary modalities are considered.

Language and communication data. Spoken or written language provides valuable insights into cognitive and emotional processes. Speech recordings, chat interactions, or narrative tasks can be analyzed to extract linguistic features such as lexical diversity, syntactic complexity, sentiment, and semantic coherence. These features may reflect emotional regulation, attention patterns, or impulsive communication behaviors.

Sleep-related physiological signals. Sleep plays a central role in emotional regulation and impulse control. Wearable devices and non-invasive sensors enable continuous monitoring of sleep parameters such as duration, efficiency, movement patterns, and cardiorespiratory activity, providing indicators of sleep quality and circadian stability.

Stress and physiological regulation. Physiological indicators of stress, including heart rate variability, skin conductance, and activity levels, provide information about emotional arousal and behavioral regulation. These signals can be obtained through wearable sensors integrated in smartwatches or other mobile health devices.

Combining these complementary modalities allows a more comprehensive representation of behavioral dynamics than single-modality approaches.

3.2. Feature Extraction

In the second component, raw sensor and interaction data are transformed into structured behavioral features using signal-processing and data-preprocessing techniques. All features described below are grounded in digital biomarker

Table 1. Maturity status of each ImpulseAI pipeline component

Pipeline component	Conceptual design	Empirical grounding / prior evidence
1 - Data acquisition		
Language speech NLP modality	Narrative tasks, chat interactions, and speech samples were collected via a mobile application over 4–8 days.	Study design and recruitment protocol piloted in [12]; NLP-based mental health indicators reviewed in [6].
Sleep monitoring Wearable sensors	Continuous home-based sleep tracking via wearable and non-invasive sensors.	Contactless cardiorespiratory sleep monitoring in home environments demonstrated in [5].
Physiological signals Smartwatch / HRV	Continuous HRV, skin conductance, and activity tracking via consumer smartwatches.	HRV as a stress and dysregulation marker is well established; consumer wearable validity reviewed in [5].
2 - Feature extraction		
Linguistic features	Lexical diversity, sentiment polarity, discourse coherence, and response latency via NLP preprocessing.	Linguistic markers of impulsivity and emotional dysregulation are demonstrated in ADHD/ICD literature, as surveyed in [6].
Sleep features	Sleep duration, efficiency, fragmentation, onset latency, and nocturnal movement indices.	Individual metrics validated as correlates of impulsivity in adolescent populations [7]; Table 2 of this paper summarizes their clinical relevance.
Physiological features HRV / autonomic	SDNN, RMSSD, LF/HF ratio, activity rhythms, and autonomic regulation markers.	HRV indices as biomarkers of emotional and impulse regulation are established in clinical literature [8].
3 - Multimodal fusion machine learning		
Fusion architecture	Feature-level fusion as primary strategy; decision-level fusion as comparison condition; fusion layer specified in pilot protocol.	Multimodal fusion for childhood mental health biomarkers was explored in [10] and [11]; the architectural choices were informed by these findings.
ML model training	Regularized classifiers, random forests, and neural architectures; cross-validation and hold-out testing protocols defined.	ML-based impulsivity classification approach piloted in predecessor study [12]; model selection criteria adapted from that work.
4 - Behavioral risk assessment		
Explainable AI (XAI)	SHAP values, feature importance analysis, and attention-based visualization are planned for the clinical output layer.	XAI methods for clinical AI transparency are well established; their integration into digital mental health is discussed in [10] and [11].
Risk profiling Decision-support output	Graded risk score with modality-specific contributions; report format designed for clinician, educator, and caregiver use.	The risk profiling concept, derived from digital biomarker frameworks [4] [10] and World Heatlh Organization guidance on AI decision support [13], informs the scope limitations.

and pediatric mental health literature and are treated as candidate behavioral indicators rather than diagnostic markers. For speech and text data, features are to be analyzed on several levels. At the vocabulary level, the framework considers diverse measures, such as the type-token ratio (TTR) and its moving-average version (MATTR). These measures give reliable insights into vocabulary richness, regardless of text length. It also incorporates the Linguistic Inquiry and Word Count (LIWC) framework, which links word use to validated psychological categories, such as emotions, social cognition, self-control, and cognitive complexity. It is used to examine psycholinguistic features. These aspects are important for understanding emotional regulation and impulsive behavior, as supported by prior studies using natural language processing in mental health research. Transformer-based language models, such as BERT and its variants, can produce contextual sentence embeddings. These embeddings help with various tasks, such as classifying sentiment, scoring semantic coherence in conversations, and identifying discourse issues. The methodology aims to analyze speech recordings and chat logs to gather data on response latency, turn-taking patterns, and interruption frequency to understand conversational dynamics. Sleep data are organized according to the American Academy of Sleep Medicine (AASM) guidelines, which define sleep stages: awake, N1, N2, N3 (slow-wave sleep), and REM. Key sleep metrics considered in the framework include total sleep time, sleep efficiency, sleep-onset latency, time spent awake after falling asleep, REM sleep latency, and REM percentage. Sleep fragmentation is assessed through the arousal index and the number of awakenings per hour. The Sleep Regularity Index (SRI) and the interdaily stability measure assess consistency in sleep timing, which is related to impulse control. Heart rate variability (HRV) should

be analyzed using time-domain measures such as SDNN (overall HRV), RMSSD (parasympathetic activity), and pNN50 (the proportion of RR-interval differences ¿50 ms). Wearable sensors may track skin conductance to measure sympathetic arousal, while accelerometers gather daily activity data, including step counts and sedentary durations.

Features across all modalities are standardized, checked for data quality, and documented with missing-data rates before being combined. The resulting feature sets comprise candidate *digital behavioral biomarkers* spanning linguistic, sleep-related, and physiological dimensions. Each feature is based on previous empirical research and chosen for its relevance to behavioral regulation and impulsivity.

3.3. Multimodal Data Fusion and Machine Learning

The third component combines the extracted features into a unified predictive model for assessing impulsivity-related behavioral risks. This pilot study considers two complementary fusion strategies: Decision-level fusion as the primary strategy and Feature-level fusion as a comparative condition. Each strategy serves a specific methodological purpose, allowing for a systematic comparison of multimodal integration approaches and fits the characteristics of the pilot dataset. Under the decision-level fusion strategy, a separate classifier is trained independently for each modality: one for linguistic features, one for sleep features, and another for physiological heart rate variability (HRV) and activity features. The predictions from each modality, represented as calibrated probability scores, are then combined using a lightweight meta-learner, specifically a logistic regression aggregator, which is trained on the outputs from the held-out validation fold. This architecture has three key advantages for the pilot context. First, each unimodal classifier operates in a lower-dimensional feature space, reducing the risk of overfitting compared to combining all features. Second, its modular structure handles missing data effectively; if sleep sensor data is absent, other modality scores can still inform the final risk estimate without needing imputation. Third, the probability outputs from each modality are interpretable, enabling a clear assessment of each modality's contribution to the overall risk score, which is essential in clinical and educational settings. In contrast, the Feature-level fusion approach concatenates the linguistic, sleep, and physiological feature vectors into a single representation after standardization and trains a single classifier on the combined input. Dimensionality reduction is applied before model training: a combination of univariate feature selection, retaining features with the strongest association with the clinical reference measure, and principal component analysis is used to reduce the joint feature space to a manageable number of components. This condition will reveal whether the added complexity of joint feature learning is justified relative to the simpler modular architecture, and the comparison between the two fusion strategies constitutes a secondary outcome of the pilot study. Before evaluating fusion conditions, each modality-specific classifier should be assessed independently. These unimodal baselines serve two purposes: they quantify the incremental predictive value of multimodal integration over a single data source, and they identify which modality contributes most strongly to risk classification. For both fusion conditions, SHAP (Shapley Additive Explanations) values should be calculated to assess the contributions of individual features to risk predictions. This approach provides modality-specific explanations that can be presented to clinicians, psychologists, or educators as part of a structured behavioral report. This design ensures that the model's outputs remain auditable and contextually relevant, aligning with the decision-support goals of the ImpulseAI framework.

3.4. Behavioral Risk Assessment

The final component focuses on identifying behavioral patterns associated with increased risk of impulse control difficulties. The integrated multimodal model generates behavioral profiles or risk indicators that support early screening and monitoring. The system is not intended to provide clinical diagnoses but to function as a decision-support tool highlighting potential behavioral risks. Such insights may support further evaluation by clinicians, psychologists, or educators and enable early preventive interventions in pediatric mental health care. Overall, the *ImpulseAI* framework demonstrates how multimodal artificial intelligence and digital biomarkers can be combined to support scalable, low-cost, and continuous behavioral monitoring in everyday environments.

4. Study design

To validate the potential of the *ImpulseAI* framework, we propose a study focusing on a multimodal, noninvasive approach to identify digital biomarkers of impulsivity in a pediatric population.

4.1. Primary Objectives and Evaluation

In the study, we will begin by evaluating the effectiveness of data acquisition techniques and the technical reliability of the sensors utilized within naturalistic home environments. Next, we will focus on identifying digital biomarkers by determining which combination of linguistic, sleep, and physiological features most effectively predicts the risk of impulsive behavior. To further our research, we will develop a behavioral profile for each participant to serve as a decision-support tool for early intervention. By adhering to this design, the study aims to create a scalable, low-cost screening tool that alleviates subjectivity and time constraints often encountered in traditional diagnostic methods.

4.2. Study Population and Recruitment

The study is designed to engage approximately 50 children and adolescents, ensuring an adequate sample size to facilitate thorough feasibility testing. Participants will be strategically recruited from various clinical and educational settings where early indicators of Impulse Control Disorders (ICDs), Attention-Deficit/Hyperactivity Disorder (ADHD), or conduct disorders are commonly identified. The participants will be strictly pediatric, spanning a broad age range to reflect developmental differences. We will focus on creating a diverse cohort in terms of emotional regulation capacities and behavioral histories, as these factors can significantly influence study outcomes.

4.3. Data Acquisition

The study will employ a comprehensive integration of mobile and wearable technologies to collect three critical data streams over 4 to 8 days. This approach is designed to identify and capture nuanced behavioral changes in participants as they navigate their everyday environments. To begin with, the first focus will be on language and communication. Participants will engage in a series of narrative tasks designed to prompt reflective storytelling and personal expression. Additionally, they will provide access to their chat interactions and recorded speech samples. This dual methodology enables researchers to analyze a wide range of linguistic features, including lexical diversity (the range of vocabulary used), syntactic complexity (the intricacy of sentence structures), and sentiment analysis (the emotional tones conveyed in the language). By gathering this data, the research aims to uncover patterns and shifts in communication that may correlate with the subjects' emotional and psychological states. The second aspect of the study will emphasize the importance of sleep monitoring, using advanced wearable devices and non-invasive sensors to collect extensive data. These devices will track vital metrics such as total sleep duration, sleep efficiency (the ratio of time spent asleep to time spent in bed), movement patterns throughout the night, and cardiorespiratory activity, including breathing rate and heart rate during sleep. This component is particularly significant, as disrupted or insufficient sleep has been tightly linked to heightened emotional reactivity and impulsive decision-making. By observing these sleep parameters, researchers can gain insights into how sleep quality may influence daily behavior and emotional stability. Finally, the study will also focus on monitoring physiological stress signals using advanced smartwatches. These devices will continuously track heart rate variability, i.e., a key indicator of the stress response, along with skin conductance, which reflects changes in sweating during emotional arousal, and overall activity levels throughout the day. By examining these physiological measurements, researchers aim to assess participants' emotional arousal and autonomic nervous system functioning, thereby gaining a deeper understanding of their psychological states and responses to their environments.

5. AI Analysis

The AI analysis aims to transform multimodal behavioral and physiological data into interpretable indicators of impulse-related risk. For validation, the model output must be compared with established clinical and behavioral assessment procedures. Standardized questionnaires completed by clinicians, teachers, and caregivers provide structured information on behavioral regulation, impulsivity, emotional stability, and developmental context. Clinical interviews complement these measures by enabling professional assessment of everyday functioning, situational triggers, developmental history, and possible comorbidities. Together, questionnaires and interviews provide a baseline against which AI-based risk indicators can be evaluated. The objective is not to replace clinical diagnosis, but to assess whether multimodal AI can support earlier screening and more targeted follow-up assessment. The processing pipeline begins

with the preprocessing and structuring of raw data from language tasks, sleep monitoring, and physiological sensors. Linguistic data are analyzed to extract indicators such as lexical diversity, discourse coherence, sentiment, response latency, and speech dynamics. Sleep data are processed to derive metrics such as sleep duration, sleep efficiency, sleep fragmentation, sleep onset latency, nocturnal movement, and irregularity of sleep timing. Physiological data are analyzed to extract stress-related indicators, including heart rate variability, activity patterns, and markers of autonomic regulation. These features are selected because they may reflect different aspects of behavioral regulation, emotional arousal, and cognitive control. In the next step, the extracted features are synchronized and integrated across modalities. This step is essential because relevant behavioral patterns may emerge from interactions among language, sleep, and physiological regulation rather than from any single modality alone. For example, reduced sleep quality may influence emotional reactivity, while stress-related physiological changes may affect speech rhythm or communication behavior. Multimodal fusion can be implemented at the feature level, by combining extracted indicators into a shared representation, or at the decision level, by aggregating predictions from modality-specific models. Machine learning models are then trained to identify patterns associated with impulsivity-related behavioral risk. Depending on the available data volume and quality, suitable models may include regularized classifiers, random forests, gradient boosting methods, support vector machines, or neural network architectures. Model development should include cross-validation, hold-out testing, and careful control of overfitting, particularly because pediatric datasets are usually limited in size and heterogeneous. Performance should be evaluated using appropriate metrics, such as sensitivity, specificity, balanced accuracy, F1 Score, and the area under the receiver operating characteristic curve. In this context, sensitivity is especially important because the framework is intended for early screening and should minimize missed cases of risk. A further step concerns the calibration and interpretation of the resulting risk indicators. Instead of producing a binary diagnostic label, the system should generate a graded behavioral risk profile that summarizes relevant findings across modalities. This profile may include modality-specific contributions, such as linguistic irregularities, sleep disturbances, or stress-related physiological patterns. Such an output can help clinicians, psychologists, or educators decide whether further assessment is needed. Finally, explainable artificial intelligence methods are applied to improve transparency and clinical usability. Techniques such as Shapley additive explanations, feature importance analysis, and attention-based visualization can help identify which linguistic, sleep-related, or physiological indicators contribute most strongly to a predicted risk score. This is particularly important in pediatric mental health, where model outputs must be understandable, cautious, and clinically interpretable. The resulting AI output is therefore intended as a decision-support instrument for early screening and preventive intervention, not as an autonomous diagnostic system.

6. Discussion and conclusion

The *ImpulseAI* framework supports the early identification of Impulse Control Disorders (ICDs) in children and teens by converting behavioral and physical signals into clear digital markers, creating an accessible mental health screening tool [4, 10, 11]. By leveraging wearable devices, the proposed framework identifies subtle behavioral changes that traditional methods may miss, supporting earlier clinical assessment and reducing delays in referral and intervention planning [4, 10]. This system supports continuous screening and combines speech patterns and sleep habits to uncover complex links between thoughts, feelings, and bodily responses [10, 11]. Additionally, using Explainable AI techniques, it delivers understandable outputs for doctors and educators, enabling informed decision-making without relying on opaque technology [6, 11]. Digital biomarkers should be interpreted as supportive indicators rather than stand-alone diagnostic evidence, strengthening rather than replacing professional clinical judgment [13]. Traditional clinical assessments, which include interviews, caregiver questionnaires, and various rating scales, typically rely on the subjective recall of experiences and behaviors [3, 14, 15]. These methods often involve interpretation and inference by the clinician based on information provided by observers. In contrast, a digital biomarker offers a more objective and precise approach to measuring behaviors [4, 10]. It captures real-time data on specific actions or physiological responses that may indicate ICD [4, 10]. This data is time-stamped and collected within a structured or semi-structured setting, allowing for a clearer and more accurate assessment of the behavior in question [4]. By using digital biomarkers, clinicians can gain insights into behaviors as they occur, providing a more reliable basis for evaluation and intervention than traditional methods that rely heavily on personal interpretation [4, 10, 14]. In a multimodal approach, different non-invasive tools capture distinct facets of impulsivity, providing a quantifiable

Table 2. Exemplary sleep-related indicators and their potential relevance for impulse control disorders

Sleep Metric	What It Reflects	ICD Relevance
Reduced total sleep time	Insufficient restorative Slow-Wave sleep	Lowered threshold for impulsive aggression next day.
Low sleep efficiency	Fragmented, non-restorative sleep	Poor overnight emotional memory processing leading to rebound irritability.
Delayed circadian phase	Circadian misalignment	Increased evening impulsivity (risk for oppositional, destructive acts).
Reduced REM latency	Dysregulated cholinergic/ aminergic balance	Correlates with poor mood regulation and reward sensitivity (kleptomania, pathological gaming).

metric [10, 11]. In that sense, measuring children's behavior can be done with a variety of innovative digital tools that provide a more quantifiable approach than traditional subjective assessments [4, 10].

Many ICDs often remain undiagnosed for extended periods, primarily because children may hide their behaviors or parents might normalize them [14, 15]. However, sleep data collected from consumer wearables like Fitbits or Garmin devices can serve as an early warning sign [5, 7]. A concerning pattern emerges when a child exhibits persistently unexplained sleep duration, frequent awakenings after sleep onset, and significant irregularity in daytime activity levels, indicating restlessness during tasks [7]. While this pattern alone does not guarantee a diagnosis of ICD, it does suggest a degree of neurophysiological dysregulation that warrants a structured assessment of impulse control [1, 2, 7, 16]. Conversely, a normal sleep profile could indicate that the underlying issues might be at play [7]. Also, by incorporating sleep measurements, the proposed framework offers valuable points for analysis and comparison for a more comprehensive diagnosis of patients [5, 7, 10]. For instance, it could help identify isolated explosive outbursts linked to specific sleep deficits or detect correlations when sleep disturbances are chronic [7, 16].

Sleep data provide unique insights into the physiological and behavioral factors that contribute to dysregulated arousal, poor impulse control, and vulnerabilities in circadian rhythms associated with these conditions [7, 16]. It is well established that impulse control is closely tied to the maturity of the prefrontal cortex and emotional regulation [1, 2, 14, 16]. Slow-wave sleep and REM sleep are essential for restoring prefrontal cortex function and regulating the amygdala, the brain region involved in emotional responses [7, 16]. Consequently, chronic sleep disruption can mimic frontal lobe dysfunction, leading to increased reactivity, reduced inhibition, and heightened reward-seeking behaviors [7, 16]. Therefore, sleep data can serve as an early, objective biomarker of the neurobiological factors that predispose individuals to impulsive actions when compromised [5, 7, 16]. In this context, sleep data can help determine whether explosive outbursts are related to transient sleep deficits, chronic dysregulation, or broader behavioral patterns requiring further clinical evaluation [7, 10, 16]. Table 2 summarizes exemplary sleep-related indicators and their potential relevance for impulse-related behavioral patterns.

The conceptual architecture of ImpulseAI is grounded in a prior empirical effort conducted within the same research group. Scherz et al. [12] addressed the problem of diagnosing impulse control disorders in youth through a scalable study design that combined standardized questionnaires, specialist clinical evaluation, mobile sensor data collection, and speech analysis in a pediatric population. This Framework represents the methodological next step: whereas Scherz et al. validated the data collection pipeline and demonstrated that machine learning classification of impulsivity-related patterns is tractable in this population, the ImpulseAI framework extends this foundation in three directions: First, it improves how we gather features by using specific, clinically relevant methods for each data type. This includes AASM sleep staging, psycholinguistic analysis using LIWC, and standardized heart rate variability measures. Second, it creates a clear structure for combining different types of data, including both decision-level and feature-level fusion, which were not detailed in the earlier study. Third, it applies explainable AI techniques to produce interpretable risk profiles rather than unclear classification outputs. This ensures transparency, which is crucial for use in pediatric clinical and educational settings. Together, these extensions constitute a progression from a validated feasibility study toward a deployable decision-support framework, with the forthcoming pilot study designed to evaluate the full integrated pipeline under the conditions established by the prior work.

7. Conclusion

The *ImpulseAI* framework provides a scalable, cost-effective screening tool for pediatric Impulse Control Disorders (ICDs) suitable for use in schools and homes. This proposal addresses a gap in pediatric multimodal solutions, as most research focuses on single modalities or on adults. The framework serves as a decision-support tool to facilitate earlier preventive interventions and improve long-term outcomes for children and adolescents. The next step involves a pilot study with about 50 participants to assess the framework's feasibility in real-world settings, focusing on data acquisition, feature extraction, and multimodal classification to identify impulsivity patterns. This will help evaluate its practical application in clinical and educational decision-making.

Acknowledgements

This research is supported by the German Federal Ministry of Education and Research (BMBF) within the framework of the program *Wissenschaftlich-Technologische Zusammenarbeit (WTZ)* with Colombia (no. 01DN24016).

References

[1] Samuel R. Chamberlain and Barbara J. Sahakian. The neuropsychiatry of impulsivity. *Current Opinion in Psychiatry*, 20(3):255–261, 2007.
[2] Jeffrey W. Dalley and Trevor W. Robbins. Fractionating impulsivity: Neuropsychiatric implications. *Nature Reviews Neuroscience*, 18(3):158–171, 2017.
[3] Jon E. Grant and Samuel R. Chamberlain. Impulsive action and impulsive choice across substance and behavioral addictions. *Addictive Behaviors*, 39(11):1632–1639, 2014.
[4] Ellen W. McGinnis, Josh Cherian, and Ryan S. McGinnis. The state of digital biomarkers in mental health. *Digital Biomarkers*, 8:210–217, 2024.
[5] Andrei Boiko, Natividad Martínez Madrid, and Ralf Seepold. Contactless technologies, sensors, and systems for cardiac and respiratory measurement during sleep: A systematic review. *Sensors*, 23(11):5038, 2023.
[6] Brittany Wolff et al. Artificial intelligence and natural language processing in clinical neuropsychology: A narrative review. *The Clinical Neuropsychologist*, 2025.
[7] Sara Uccella, Ramona Cordani, Federico Salfi, et al. Sleep deprivation and insomnia in adolescence: Implications for mental health. *Brain Sciences*, 13(4):569, 2023.
[8] Karolina Kozak, Aliya M. Lucatch, Darby J. E. Lowe, Iris M. Balodis, James MacKillop, and Tony P. George. The neurobiology of impulsivity and substance use disorders: implications for treatment. *Annals of the New York Academy of Sciences*, 1451(1):71–91, 2019.
[9] Gregory T. Smith and Melissa A. Cyders. Integrating affect and impulsivity: The role of positive and negative urgency in substance use risk. *Drug and Alcohol Dependence*, 163:S3–S12, 2016.
[10] Myounglee Choo, Doeun Park, et al. Exploring a multimodal approach for utilizing digital biomarkers for childhood mental health screening. *Frontiers in Psychiatry*, 15:1348319, 2024.
[11] Zifan Jiang, Salman Seyedi, Ahmed Abbasi, et al. Multimodal mental health digital biomarker analysis from remote interviews using facial, vocal, linguistic, and cardiovascular patterns. *IEEE Journal of Biomedical and Health Informatics*, 28(3), 2024.
[12] Wilhelm Daniel Scherz, Maksym Gaiduk, Jorge Avila-Campos, Ralf Seepold, Natividad Martínez Madrid, Paula Herrera, and Julian David Echeverry. Applying machine learning to diagnose impulse control disorders in youth: Towards a scalable study design. *Procedia Computer Science*, 270:5716–5724, 2025.
[13] World Health Organization. *WHO Guideline: Recommendations on Digital Interventions for Health System Strengthening*. World Health Organization, Geneva, 2019.
[14] Russell A. Barkley. Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of adhd. *Psychological Bulletin*, 121(1):65–94, 1997.
[15] American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5-TR*. American Psychiatric Association Publishing, Washington, DC, 5 edition, 2022.
[16] B. J. Casey, Rebecca M. Jones, and Todd A. Hare. The adolescent brain. *Annals of the New York Academy of Sciences*, 1124(1):111–126, 2008.