

Mobile, Wearable, and Contactless Approaches to Pediatric Stress Detection: A PRISMA-ScR Guided Scoping Review

Arshad Nasser, Sami Elferik, Malak Salim Omar Baslyman

Submitted to: JMIR mHealth and uHealth
on: September 29, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 49

 Multimedia Appendixes 50

 Multimedia Appendix 1..... 50

 Multimedia Appendix 2..... 50

 Multimedia Appendix 3..... 50

Mobile, Wearable, and Contactless Approaches to Pediatric Stress Detection: A PRISMA-ScR Guided Scoping Review

Arshad Nasser¹ BTECH, MDes, PhD; Sami Elferik¹; Malak Salim Omar Baslyman¹

¹ King Fahd University of Petroleum and Minerals Dhahran SA

Corresponding Author:

Arshad Nasser BTECH, MDes, PhD

King Fahd University of Petroleum and Minerals
Academic Belt Road
Dhahran
SA

Abstract

Background: Stress in childhood and adolescence can disrupt emotional, cognitive, and physiological development, with long-term health consequences. Conventional assessments (eg, cortisol sampling, retrospective questionnaires) are invasive, burdensome, and poorly suited for pediatric use. Advances in mobile health (mHealth), wearable devices, and non-contact sensing enable non-invasive, real-time monitoring of stress in children and adolescents.

Objective: This scoping review synthesizes evidence on non-invasive technologies for pediatric stress detection and regulation, with particular attention to mobile, wearable, and ubiquitous health applications. It maps the technologies, protocols, and interventions reported to date and identifies gaps and future directions for research and practice.

Methods: Following PRISMA-ScR guidelines, five databases (PubMed, Scopus, IEEE Xplore, ACM Digital Library, Google Scholar) were searched for peer-reviewed studies published between 2013 and June 2025. Eligible studies involved participants aged 0–18 years, used non-invasive technological methods for stress detection, monitoring, or mitigation, and reported empirical outcomes. Data were charted across seven dimensions: technology type, stress induction and validation, study context, age group, evaluation method, datasets, and intervention strategies.

Results: Forty-nine studies were included. Most used wearable devices (84%), with limited exploration of non-contact (14%) or hybrid systems. Stress induction methods were heterogeneous, dominated by cognitive and academic tasks, with no standardized pediatric protocols. Adolescents were the most studied group, whereas children under 8 years were underrepresented. Evaluation approaches ranged from traditional statistics to supervised and deep learning models (accuracies 82%–95.6%), though reproducibility was constrained by small, homogeneous samples and a lack of open pediatric datasets. Only 9 studies integrated mitigation strategies (eg, biofeedback, serious games, caregiver-facing tools), typically in short-term deployments. Ethical and privacy considerations were acknowledged but rarely operationalized. While the review covers a broad spectrum of non-invasive approaches, the majority of included studies were delivered through mobile and wearable platforms, with fewer exploring ambient or ubiquitous sensing, aligning the field closely with mobile and ubiquitous health research.

Conclusions: Non-invasive mobile and wearable technologies offer strong potential for scalable pediatric stress monitoring, yet the field remains fragmented. Priorities include developing age-appropriate stress protocols, creating diverse open datasets, evaluating systems in real-world and longitudinal contexts, and embedding closed-loop interventions. Stronger ethical, privacy, and equity safeguards are critical for translation into schools, telehealth, and everyday mHealth practice. Clinical Trial: Not Applicable

(JMIR Preprints 29/09/2025:84951)

DOI: <https://doi.org/10.2196/preprints.84951>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ Please make my preprint PDF available to anyone at any time (recommended).

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in [http://www.jmir.org/2018/12/e12345/](#), I will be able to make my manuscript PDF available to all users.

✓ **No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in [http://www.jmir.org/2018/12/e12345/](#), I will be able to make my manuscript PDF available to all users.**



Original Manuscript

Mobile, Wearable, and Contactless Approaches to Pediatric Stress Detection: A PRISMA-ScR Guided Scoping Review

Arshad Nasser
Sami Elferik
Malak Basyman

Background: Stress in childhood and adolescence can disrupt emotional, cognitive, and physiological development, with long-term health consequences. Conventional assessments (eg, cortisol sampling, retrospective questionnaires) are invasive, burdensome, and poorly suited for pediatric use. Advances in mobile health (mHealth), wearable devices, and non-contact sensing enable non-invasive, real-time monitoring of stress in children and adolescents.

Objective: This scoping review synthesizes evidence on non-invasive technologies for pediatric stress detection and regulation, with particular attention to mobile, wearable, and ubiquitous health applications. It maps the technologies, protocols, and interventions reported to date and identifies gaps and future directions for research and practice.

Results: Forty-nine studies were included. Most used wearable devices (84%), with limited exploration of non-contact (14%) or hybrid systems. Stress induction methods were heterogeneous, dominated by cognitive and academic tasks, with no standardized pediatric protocols. Adolescents were the most studied group, whereas children under 8 years were underrepresented. Evaluation approaches ranged from traditional statistics to supervised and deep learning models (accuracies 82%–95.6%), though reproducibility was constrained by small, homogeneous samples and a lack of open pediatric datasets. Only 9 studies integrated mitigation strategies (eg, biofeedback, serious games, caregiver-facing tools), typically in short-term deployments. Ethical and privacy considerations were acknowledged but rarely operationalized. While the review covers a broad spectrum of non-invasive approaches, the majority of included studies were delivered through mobile and wearable platforms, with fewer exploring ambient or ubiquitous sensing, aligning the field closely with mobile and ubiquitous health research.

Conclusions: Non-invasive mobile and wearable technologies offer strong potential for scalable pediatric stress monitoring, yet the field remains fragmented. Priorities include developing age-appropriate stress protocols, creating diverse open datasets, evaluating systems in real-world and longitudinal contexts, and embedding closed-loop interventions. Stronger ethical, privacy, and equity safeguards are critical for translation into schools, telehealth, and everyday mHealth practice.

Keywords: mHealth; wearable technology; pediatric stress; child; adolescent; non-invasive monitoring; telehealth

1. Introduction

Stress during childhood and adolescence, if unmanaged, can significantly impact psychological, emotional, and physiological development, potentially leading to chronic health issues such as anxiety disorders, cardiovascular problems, and cognitive impairments later in life [1–3]. Recognizing and addressing stress early in these formative years is therefore crucial for promoting healthy long-term outcomes.

Traditionally, stress detection and assessment have relied on invasive or laboratory-based methodologies, such as cortisol measurements from saliva or blood samples, or extensive psychometric evaluations conducted retrospectively once stressful events have passed [4,5]. Despite their validity, these methods pose logistical, ethical, and emotional challenges in pediatric populations, where invasive procedures and prolonged assessments may be uncomfortable or impractical [6].

Recent technological advances have propelled the development of non-invasive technologies for real-time, unobtrusive monitoring and mitigation of stress. In pediatric contexts, these systems are frequently deployed as mobile health (mHealth) applications or wearable devices, with a smaller body of work exploring ambient or ubiquitous sensing platforms [7,8]. Such technologies leverage physiological indicators such as heart rate variability (HRV), electrodermal activity (EDA), and respiratory patterns, as well as behavioral markers including facial expressions, vocal intonation, and movement patterns, to infer stress levels within children's daily lives [9–11]. Their integration into school routines, home environments, and telehealth platforms highlights their potential to bridge laboratory-based research with everyday health practice.

However, despite the rapid proliferation of these non-invasive technologies, the current research landscape remains fragmented and predominantly adult-focused, with significant inconsistencies in methodological rigor and evaluation standards [12]. González Ramírez et al. [12] and Arsalan et al. [13] recently highlighted these gaps, noting deficiencies in technological integration, validation methodologies, developmental appropriateness, and standardized evaluation frameworks across studies. Importantly, research explicitly targeting pediatric populations remains markedly underrepresented, creating a substantial knowledge void regarding the applicability and effectiveness of non-invasive stress detection technologies among children and adolescents, who differ considerably from adults in their physiological, psychological, and developmental profiles [14,15].

Given the distinct developmental trajectories, stressors, and ethical considerations inherent in pediatric research, there is an urgent need for a comprehensive mapping of current non-invasive technologies, the physiological and behavioral parameters they assess, validated methods for inducing and measuring stress, and contextual factors unique to children and adolescents.

Recent advances in wearable sensors, behavioral biometrics, and contactless physiological monitoring methods, including photoplethysmography (PPG), galvanic skin response (GSR), accelerometry, and microwave reflectometry, offer promising avenues for pediatric stress assessment without intrusive engagement [16]. These technologies offer distinct advantages, such as minimizing emotional disruption, enhancing compliance and comfort, enabling continuous and context-aware monitoring, and addressing

ethical and practical concerns by avoiding physical intrusion and procedural anxiety [17–20].

This scoping review systematically addresses these gaps by synthesizing literature focused explicitly on non-invasive technologies designed for stress detection, monitoring, and intervention in pediatric populations. While diverse in form, the majority of such systems are implemented through mobile and wearable platforms, with fewer examples of ambient or ubiquitous sensing. Specifically, this review aims to: (1) identify the technologies and physiological or behavioral parameters assessed; (2) examine validated methodologies for reliably inducing and measuring stress; (3) explore contexts and developmental stages where these technologies have been deployed; (4) evaluate frameworks and metrics used for assessing technological accuracy and efficacy; (5) review available datasets, including their characteristics and collection methodologies; (6) assess intervention strategies integrated into these systems for stress regulation; and (7) identify limitations and gaps related to long-term monitoring, inclusivity, cultural considerations, and practical implementation. Table 1 summarizes the seven research questions that guided this review, grouped under four thematic categories: Detection, Evaluation, Mitigation, and Cross-cutting challenges.

Table 1. Categorized research questions for the scoping review

Category	Research Question	Purpose / Focus
Detection	RQ1. What non-invasive technologies and modalities have been used to detect stress in children and adolescents, and what parameters do they assess?	Identify tech types (wearables, apps, etc.) and the stress-related signals measured
	RQ2. How is stress induced and validated or evaluated in children/pediatric studies?	Examine stress induction protocols (e.g., SCWT, IAPS) and ground truth methods
	RQ3. In what settings and age groups have these technologies been applied?	Explore deployment contexts (home, school, clinic) and developmental stages
Evaluation	RQ4. What detection methods, frameworks and evaluation approaches are used, and how accurate or effective are they?	Covers ML/heuristic models, detection frameworks, and their performance metrics, Explore qualitative, quantitative, and hybrid evaluation methods
	RQ5. What datasets are available for stress detection/evaluation in children, and what are their characteristics and collection methodologies?	Summarize dataset properties: size, modality, annotation, accessibility
Mitigation	RQ6. What mitigation strategies or intervention frameworks have been integrated	Include biofeedback, adaptive interventions, and real-time strategies

Category	Research Question	Purpose / Focus
	with these systems to reduce stress in children?	
Cross-cutting	(Discussion theme) What are the limitations in addressing long-term stress monitoring, inclusivity, and cultural diversity in the current literature?	Highlight gaps in real-world applicability, ethical fit, and population coverage

By addressing these research questions, this scoping review not only synthesizes current scientific knowledge but also establishes a foundational framework for future research and practical applications, ultimately aiming to enhance stress management strategies and improve pediatric health outcome.

2. Methodology

2.1. Review Framework

This scoping review was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) [66]. We selected a scoping review approach as it enables systematic mapping of existing research literature, Mobile, Wearable, and Contactless Approaches to Pediatric Stress Detection: A PRISMA-ScR Guided Scoping Review 3 identification of key concepts, and clarification of research gaps without appraising methodological quality, which aligns with the exploratory objectives of our review [45].

2.2. Search Strategy

We searched four bibliographic databases (PubMed, Scopus, IEEE Xplore, and ACM Digital Library) and screened Google Scholar as a supplementary source. Searches covered January 2013 to June 2025 (last search: June 2025). Database-specific strategies and exact queries are in Multimedia Appendix 2.

- **Population:** "children," "adolescents," "pediatric," "youth."
- **Technology:** "non-invasive," "wearable," "ambient sensors," "mobile applications," "contactless," "physiological monitoring," "behavioral biometrics", "ubiquitous computing."
- **Outcome:** "emotion," "anxiety," "stress," "stress detection," "stress monitoring," "stress regulation."

The search queries were adapted according to each database's syntax and specifications, limited to articles published in English from January 2013 to June 2025.

2.3. Inclusion and Exclusion Criteria

Studies eligible for inclusion met the following criteria:

- Focused explicitly on children and adolescents (0–18 years).
- Evaluated non-invasive technological methods for stress detection, monitoring, or mitigation.
- Reported empirical outcomes related to stress measurement or management.
- Published as peer-reviewed journal articles or conference proceedings.

Studies were excluded if they:

- Targeted adult populations exclusively.
- Utilized invasive or physically intrusive methodologies.
- Were purely theoretical, methodological, or lacked validation of any sort.
- Were literature reviews, editorial commentaries, or opinion articles.

2.4. Study Selection Process

The study selection process followed three structured phases: identification, screening, and eligibility assessment. Search results were exported into Zotero to remove duplicates and facilitate the screening process. After the elimination of duplicates, the screening stage was performed. This consisted of checking that the title and abstract of the selected studies answered the following questions about the inclusion criteria addressed: (1) Is the study related to stress? (2) Is this study related to child/children? (3) Does the study use non-invasive techniques or technologies? If all the answers were affirmative, the study was assessed for eligibility, which consisted of reading the full text to determine if it met the inclusion criteria. Finally, data from the included studies were extracted for further analysis. The full texts of potentially eligible articles were then independently assessed by the reviewers. Studies meeting the inclusion criteria were finalized for data extraction.

2.5. Data Extraction and Analysis

We followed deductive and inductive approaches to identify relevant data concerning the studies' characteristics [23,24]. Initially, we predefined a set of data categories based on existing literature and knowledge relevant to non-invasive stress detection technologies [12,25]. Using this deductive approach, we systematically collected data about interventions, including physiological signals measured (e.g., heart rate variability, electrodermal activity), behavioral indicators (e.g., facial expressions, vocal intonation), and the form factor of devices, defined as physical characteristics such as dimensions, shapes, and placement (e.g., wristbands, smartwatches, ambient sensors) [26]. Additionally, we recorded details on hardware and software specifics, and noted interaction modalities used for presenting stress-related information (visual, auditory, tactile).

Contextual data extraction included study settings (e.g., schools, homes, clinical settings), participant demographics (age groups, gender, participant type such as students), experimental conditions (controlled vs. naturalistic environments), specific stress induction methods, and validated tools employed for measuring stress levels. Furthermore, we extracted the most significant results and conclusions reported by authors regarding the effectiveness and applicability of stress detection interventions.

2.6. Study Selection

Figure 1. PRISMA-ScR flow of study selection. Records identified (n=545); duplicates removed (n=44); records screened (n=501); records excluded (n=406); full-text reports assessed (n=95); reports excluded with reasons (n=46); studies included (n=49).

3. Results

A total of 49 studies were included in this review (Figure). (Figure 1).

3.1. Publication Characteristics

The research papers published between 2013 and 2025 shows a clear growth in output over the past five years. The highest number of publications occurred in 2022 (12 studies), followed by 2023 (9 studies) and 2024 (9 studies), reflecting increasing academic interest in non-invasive stress monitoring in pediatric populations. (Figure 2).

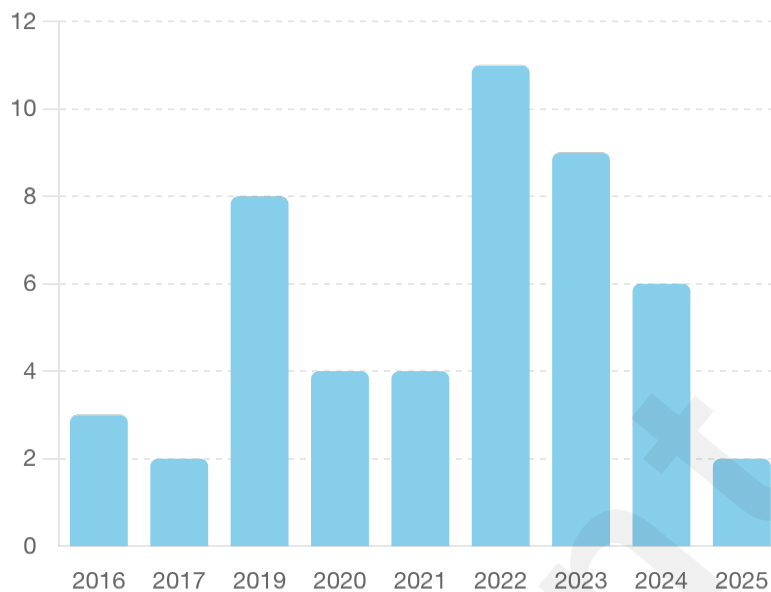


Figure 2. Distribution of included studies by year (n=49).

Regarding publication type, journal articles accounted for the largest share (29 studies, 59%), followed by conference proceedings (16 studies, 33%). A smaller number were categorized as book chapters (3 studies, 6%) and theses or preprints (1 study, 2%). This distribution illustrates both the interdisciplinary and emerging nature of the field, with a strong presence in applied computing and healthcare venues.

This diverse spread indicates the topic's cross-cutting relevance across engineering, health sciences, and educational technology disciplines.

3.2. Technologies and Modalities Used

This section categorizes the technological approaches used for stress monitoring in children and adolescents across the reviewed studies. We analyze technologies along two orthogonal axes: (1) *Deployment form factor* describes how a system is delivered (e.g., wearable wristband or smartwatch; mobile application; ambient or object-integrated device). (2) *Sensing modality* describes how signals are acquired: *contact-based physiological* (e.g., HR/HRV, EDA, temperature, ACC), *non-contact* (e.g., vision, audio, rPPG), or *hybrid*. Counts reported in Section 3.2 refer to *deployment form factor*; counts in Section 4.1 refer to *sensing modality*. Studies can appear in multiple modality buckets, so totals may not sum to $n=49$.

Table 2. Taxonomy used in this review: deployment form factor vs sensing modality. Counts in Section 3.2 refer to form factor; counts in Section 4.1 refer to sensing modality; studies may appear in multiple modality buckets.

Axis	Categories	Examples
Deployment form factor	Wearable; Mobile app;	Wristbands, smartwatches; smartphone
	Ambient/object-integrated	logging/biofeedback; classroom device/robot
Sensing modality	Contact physiological; Non-contact; Hybrid	HR/HRV, EDA, Temp, ACC; vision (RGB/thermal, rPPG), audio; voice + ACC

Technologies were grouped into three primary categories: wearable systems, ambient or object-integrated systems, and multimodal or hybrid systems. Each group is described below based on the types of devices employed, signal modalities captured, and design patterns observed. *Note: the low count of ambient/object-integrated systems does not contradict the higher non-contact tally in Section 4.1, which also includes camera- and audio-based systems that are non-contact but not embedded in the environment or objects.*

Wearable Systems

Wearable technologies constituted the majority of included systems. These included wristbands, chest straps, integrated biosensors, and mobile-linked tracking devices. Physiological signals such as heart rate (HR), electrodermal activity (EDA), heart rate variability (HRV), body temperature, and accelerometer data were the most common modalities. Several systems also captured motion or engagement metrics through mobile apps or game interactions. (Table 3).

For instance, devices like the Empatica E4 and Fitbit-style trackers were used in multiple studies to collect physiological signals such as HR, EDA, skin temperature, and motion [27–29]. Others like Antle et al.[14] embedded biosensors into a biofeedback game to track HR and EDA during user interaction. A few systems (e.g., Kadirvelu et al.[30]) collected passive behavioral and contextual data such as mood, sleep, or activity levels through mobile platforms, although physiological integration varied.

Table 3. Wearable-Based Stress Monitoring Systems

Authors (Year)	Device/System	Modality	Parameters Captured
Aktas et al. (2022) [28]	Empatica E4 sensor suite	Physiological	EDA, HR, ST, ACC
Antle et al. (2019)[14]	Wearable-integrated biofeedback game	Physio + Behavioral	HR, EDA, interaction
Buddi et al. (2019)[31]	Tracker + app system	Physio + Logs	HR, movement, usage
Chaowadee et al. (2021) [32]	Stress warning wearable unit	Physiological	HR, temperature, GSR
Choi et al. (2017)[33]	Wrist-worn biosensors	Physiological	GSR, HR, temp
Coşkun et al. (2023) [29,34,35]	Tablet-linked sensors	Physiological	EDA, HR, temperature
Deng et al. (2022)[27]	Empatica + mobile data logging	Physio + Contextual	HR, temp, EDA, motion
Kadirvelu et al. (2023)[30]	Mindcraft passive sensing app	Behavioral + Passive	Mood, sleep, motivation

Ambient and Object-Integrated Systems

Only one study deployed an ambient or embedded system. Chaowadee et al.[32] introduced a biometric alert system integrated into the classroom, providing real-time stress prompts based on HR, GSR, and temperature readings. No other systems in the final set used smart objects, robotic agents, or passive ambient sensing technologies. Despite their low representation, these systems remain relevant for contexts like schools, where non-contact, passive monitoring may be more suitable. (Table 4).

Table 4. Ambient or Object-Integrated Systems

Authors (Year)	System	Modality	Parameters Captured
Chaowadee et al. (2021) [32]	Classroom biometric alert device	Physiological	HR, Temp, GSR

Multimodal and Hybrid Systems

Multimodal systems combined physiological signals with behavioral, contextual, or self-reported inputs. In some cases, physiological sensing was supported by mobile logging (e.g., Deng et al.[27]), while others focused more on contextual awareness and affective states through passive sensing and user input (e.g., Kadirvelu et al.[30]). One system (Antle et al.[14]) combined wearable input with visual biofeedback in a serious game environment. (Table 5).

While HR and EDA were frequently used, behavioral signals varied—ranging from engagement logs to mood and sleep tracking. Signal integration and fusion strategies, however, were not described in detail in most cases.

Table 5. Multimodal / Hybrid Systems

Authors (Year)	System	Modality	Parameters Captured
Deng et al.[27] (2022)	Empatica + mobile logging	Physio + Context	HR, Temp, EDA, motion
Kadirvelu et al.[30] (2023)	Minecraft mobile sensing app	Behavior + Passive	Mood, sleep, motivation
Antle et al. (2019)[14]	Game-based biofeedback system	Physio + Visual feedback	HR, EDA

Across all modalities, HR and EDA were the most frequently captured signals. Wearable systems supported the broadest range of biosignals and were commonly integrated into mobile ecosystems. Multimodal systems enabled behavioral or contextual extensions, though few used advanced signal fusion. Ambient/object-integrated systems were underrepresented, with limited data on long-term deployment or scalability. Differences in integration depth and feedback mechanisms were noted but not systematically reported.

3.3. Stress Induction and Validation Methods

Understanding how stress is induced and validated in pediatric populations is central to evaluating the effectiveness and ecological relevance of non-invasive stress monitoring technologies. Across the 28 out of 49 studies reviewed, we observed a wide range of stress induction paradigms—both explicit and implicit—implemented in diverse settings such as classrooms, laboratories, clinics, and home environments. These paradigms varied not only in intensity and modality but also in the underlying rationale for their use, often reflecting age-appropriateness, ethical considerations, and target population characteristics.

Stress Induction Modalities

We identified four major categories of stress induction: cognitive/academic tasks, social or evaluative scenarios (including robot-mediated interactions), game-based or interactive platforms, and naturalistic or passive conditions. These categories are reflected in Tables 6–9, each summarizing studies grouped by induction approach, validation method, instrumentation, and context.

Cognitive and academic stressors, including timed quizzes, memory tasks, and exam-period assessments, were used to trigger performance-related stress through cognitive load and time pressure. These were especially prevalent in school settings, where stress is often linked to academic performance and attention demands (e.g., Buddi et al., 2019[31]; Jin et al., 2023[36]). While standardized tasks such as the Stroop test were infrequent, many studies simulated academic challenges that closely mirrored real-world classroom conditions. (Table 6).

Social and evaluative stressors were typically deployed in controlled environments and involved mock interviews, public speaking, or emotionally salient interactions with humanoid robots. These aimed to elicit socially triggered stress responses, often through perceived judgment or performance anxiety. Several studies integrated robot-mediated interactions to explore the intersection of novelty, social tension, and behavioral response (e.g., Chaowadee et al.[32]; Coşkun et al.[29]). (Table 7).

Game-based and interactive paradigms offered a controlled yet engaging environment for stress induction. Games were designed to include competitive elements, time constraints, or escalating task difficulty, making them particularly suitable for child and adolescent populations. Unlike traditional cognitive tasks, these settings allowed for richer engagement and the integration of embedded behavioral tracking (e.g., Antle et al. [14]; Vural et al.[37]). (Table 8).

Finally, a substantial subset of studies adopted naturalistic or passive paradigms, refraining from introducing any artificial stressors. Instead, they monitored stress as it emerged during daily routines, classroom activities, or home life. These methods were typically paired with ecological momentary assessments (EMA), caregiver tagging, or passive sensing through mobile and wearable platforms. Though less controlled, these studies offered high ecological validity and ethical accessibility, especially for longitudinal or clinical deployments (e.g., Kadirvelu et al.[30]; Yu et al.[38]). (Table 9).

Validation Methods

Validation approaches across studies varied in both modality and rigor. Physiological measures—such as heart rate variability (HRV), electrodermal activity (EDA), and electrocardiography (ECG)—were the most commonly employed, offering objective correlates of autonomic arousal. In a smaller subset, EEG and thermal imaging were used to capture neurophysiological or affective markers under stress.

Self-report tools were also frequently used, particularly in school-aged children and adolescents. Standardized instruments like the Perceived Stress Scale (PSS), the State-Trait Anxiety Inventory for Children (STAI-C), or customized visual scales were used to gather subjective assessments of stress. However, the appropriateness of self-report was dependent on the age and cognitive capacity of the participants, with many younger children requiring pictorial or observational alternatives.

Mixed-method approaches were increasingly common, combining physiological sensing with self-report, observational coding, or behavioral tagging. These studies often demonstrated higher validation robustness and better contextual grounding, especially in naturalistic and home-based deployments (e.g., Isbister et al., 2022[39]; Lutin et al., 2023[40]). A few studies also employed clinician assessments or video-coded behavioral markers to validate stress in populations with limited communication abilities (e.g., Redd et al., 2020[41]).

Structure of the Evidence

To enhance readability and interpretability, we present the studies in four separate tables grouped by the nature of their induction paradigm: cognitive/academic (Table 6), social or evaluative including robot-based (Table 7), game-based or interactive (Table 8), and naturalistic or passive (Table 9). These groupings reflect not only methodological variation but also differing assumptions about stress ecology, feasibility, and ethical constraints across pediatric contexts.

Table 6. Table 5. Studies Employing Cognitive and Academic Stressors

Author(s) (Year)	Induction Method	Validation Method	Signals/Instruments	Context Used
Azevedo et al. (2016) [17]	Timed memory game with performance pressure	Self-report (Likert-based)	Behavioral	Computer lab (school)
Buddi et al. (2019)[31]	Timed quizzes via mobile app	Self-report scale post-task	Physiological, Behavioral	School (in-the-wild deployment)
Zhang et al. (2020)[42]	Natural (exams, academic/social pressure)	User feedback on app: emotion score <0.5 = stress	Text, Image, Sleep & Exercise data	Home + school, over 7–36 weeks

Author(s) (Year)	Induction Method	Validation Method	Signals/Instruments	Context Used
Jin et al. (2023)[36]	Academic exam periods (naturalistic)	Self-report questionnaires (stress scales)	Physiological	School

Table 7. Studies Employing Social, Evaluative, or Robot-Based Stressors

Author(s) (Year)	Induction Method	Validation Method	Signals/Instruments	Context Used
Coşkun et al. (2023)[34]	Public speaking task	HRV, self-report	ECG, self-report scales	School
Chaowadee et al. (2021)[32]	Robot interaction in classroom	Observer ratings, HR, GSR	Video, physiological	School
Goulart et al. (2019)[43]	Scripted robot sessions with emotion triggers	Observer/emotion coders	Thermal, Visual, Behavioral	Lab
Kumazaki et al. (2017)[44]	Social robot-based conversations	Observer-coded video behavior	Behavioral	Clinic
Loftness et al. (2022)[45]	Mock interviews & tasks	Observer ratings + PSS-C	Physiological, Visual, Audio	School
Sarabadani et al. (2020)[46]	Attention/social triggers in behavior sessions	Video-coded behavioral categories	Physiological	Clinical

Table 8. Studies Using Game-Based or Interactive Stressors

Author(s) (Year)	Induction Method	Validation Method	Signals/Instruments	Context Used
Antle et al. (2019) [14]	Gameplay challenges and narrative stressors	Pre/post self- report + session logs	Physiological + Behavioral	School
Carlier et al. (2019) [47]	Mild challenges within game environment	Self-report (adapted for ASD)	Behavioral	Therapeutic center
Coşkun et al. (2023)[29]	Timed and competitive game scenarios	Observer + self- report	Physiological	Therapeutic center
Vural et al. (2024) [37]	Physiotherapy game activity	Video segment annotated by 3 therapists	Visual (facial images)	Clinic + home (serious game sessions)

Author(s) (Year)	Induction Method	Validation Method	Signals/Instruments	Context Used
		(voting)		
Kim et al. (2021) [48]	Educational games with increasing difficulty	Self-report and game performance	Physiological	School
Nicolaidou et al. (2019)[49]	Stress scenarios within game	Self-reported stress (numeric scales)	Behavioral (interactions, self-reports)	School
Theofanopoulou et al. (2019)[50]	Game-based tasks with performance pressure	Self-report + observed behavior	Behavioral, Physiological	Lab/school

Table 9. Studies Using Naturalistic or Passive Monitoring.

Author(s) (Year)	Induction Method	Validation Method	Signals/Instruments	Context Used
Lutin et al. (2023)[40]	Emotionally evocative vignettes; daily experience sampling	Self-report via stress thermometer and pictorial scales	Subjective (valence, arousal), verbal reasoning	Home / School (self-report, natural context)
Kadirvelu et al. (2023) [30]	Naturalistic (daily life)	Self-report + passive logging	Active: self-reported; Passive: step count, ambient light, app use, noise, battery, location	Home + school (real-world)
Isbister et al. (2022)[39]	Naturalistic interactions	Qualitative observations, post-interaction interviews	Behavioral, Interaction data	Lab
Masino et al. (2019)[51]	Naturalistic school routines	Teacher and therapist annotations	Physiological	School
Melinda et al. (2024) [52]	Naturalistic	Fuzzy logic model (no ground truth comparison)	Physiological	Prototype setting
Nguyen et	Daily life (in-situ)	Self-reported	Physiological	Home

Author(s) (Year)	Induction Method	Validation Method	Signals/Instruments	Context Used
al. (2021) [53]		awareness + journaling		
Northrup et al. (2016) [54]	Naturalistic (daily classroom and therapy activities)	Caregiver observation matched with sensor data	Physiological	School + therapy environments
Payani et al. (2021)[55]	Assumed daily stress (not evaluated experimentally)	None (prototype phase)	Physiological	Conceptual / Lab
Redd et al. (2020)[41]	Naturalistic (home and clinic)	Clinician rating + behavioral diaries	Physiological	Home + clinic
Wang et al. (2025)[56]	Naturalistic home setting	Parental observations + reflection	Behavioral	Home
Yao et al. (2023)[57]	Homework/learning stress (natural)	Self-report and use context	Behavioral	Home + school
Yu et al. (2023)[38]	Naturalistic	Caregiver-tagged stress episodes	Physiological + Behavioral	Home

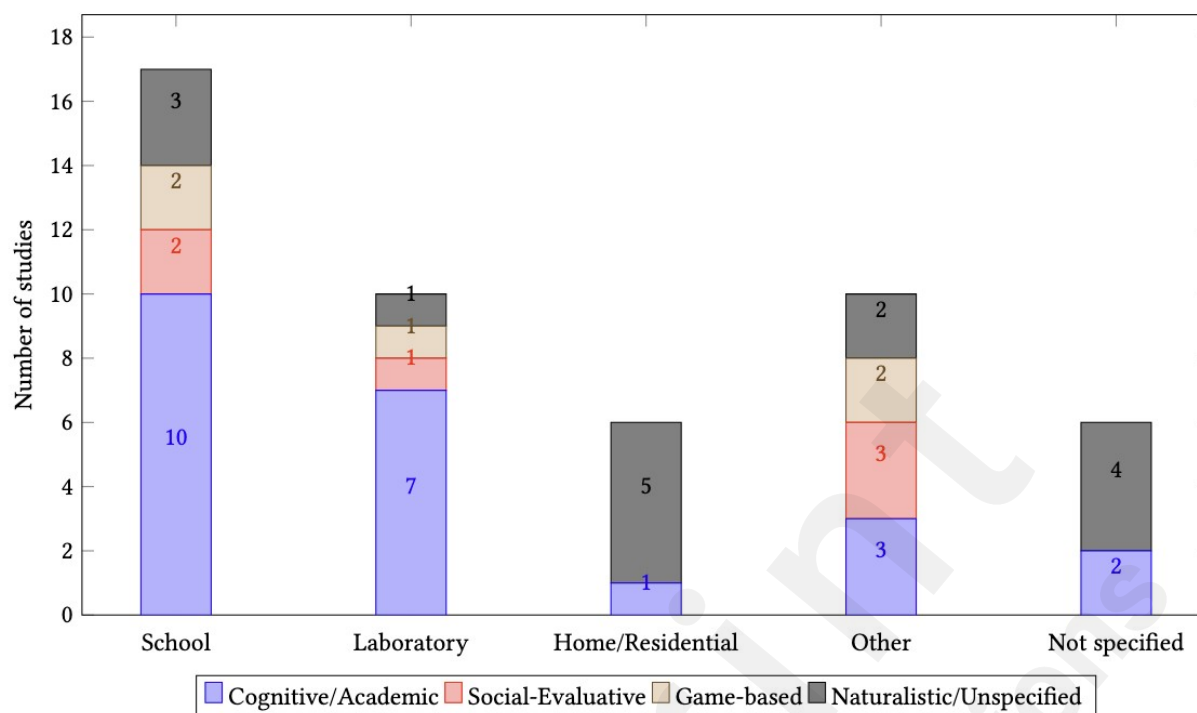


Figure 3. Distribution of study environments by stressor type (n=49). Hybrid, conceptual, or undefined stressors were collapsed into the “Not specified” category.

Notably, there was considerable variability in stress induction intensity, validation methods, and rigor across reviewed studies. Laboratory-based stressors provided consistency and controlled replication but often lacked ecological realism compared to naturalistic approaches. Validation techniques also varied significantly, with physiological methods offering greater objectivity but less subjective nuance compared to self-report measures. These differences underline the complexity and necessity of carefully selecting appropriate induction and validation methods tailored to specific pediatric contexts and research objectives.

3.4. Study Contexts and Age Groups

This section synthesizes the deployment contexts and age groups targeted by the reviewed studies. Understanding the contexts and age-specific applicability of non-invasive stress-monitoring technologies is critical for assessing their ecological validity, usability, and real-world implementation potential.

The reviewed studies spanned various environments, categorized primarily as controlled laboratory settings, school-based deployments, home or residential contexts, and hybrid approaches combining multiple contexts. Laboratory environments provided controlled conditions, ideal for standardized stress induction and initial physiological validation, as exemplified by the studies of Azevedo et al. (2016)[17] and Nicolaidou et al. (2019)[49]. However, the controlled nature of these studies potentially limited ecological validity, especially for younger children who may find such settings stressful or unfamiliar. (Table 10).

School-based studies were commonly employed, providing significant ecological validity due to their direct relevance to daily stressors experienced by children and adolescents, such as academic pressures and peer

interactions. Examples include Buddi et al. (2019)[31], who utilized a mobile app integrated with school routines, and Antle et al. (2019)[14], who embedded biofeedback technologies within school settings. Despite their contextual strength, school-based studies seldom tracked stress management interventions longitudinally.

Home or residential settings were underrepresented despite their potential to reflect everyday stressors and natural behavioral responses. Studies such as those by Lutin et al. (2023)[40] uniquely combined home and school contexts, capturing richer ecological insights. Still, such cross-context integration was rare.

Hybrid or mixed-context studies provided a balanced perspective by blending controlled evaluations with naturalistic monitoring, exemplified by the approach of Deng et al. (2022)[27], who combined passive sensing at home and school. These designs offer nuanced insights into stress responses across varied environments but remain relatively uncommon.

Targeted Age Groups

Across the 49 included studies, adolescents (13–18 years) and children under 13 years were equally represented (18 studies each, 36.7%). Nine studies (18.4%) spanned mixed-age pediatric samples, while four studies (8.2%) did not clearly report participant age. Notably, children younger than 8 years were markedly underrepresented. (Table 11).

Studies predominantly targeted older children and adolescents, generally aged 8–14 years, reflecting developmental stages where self-reporting and adherence to sensor use are more feasible. For instance, Buddi et al. (2019)[31] and Aktas et al. (2022)[28] targeted pre-adolescents and adolescents, typically within school or therapeutic contexts. Adolescents often constituted the focus due to their increased capacity for reliably engaging with sensor-based technologies and articulating subjective stress experiences.

Methodological and ethical challenges in younger cohorts—such as limited verbalization of stress, sensitivity to sensor ergonomics, and heightened ethical considerations—likely contributed to their underrepresentation. Nonetheless, some studies, such as that by Lutin et al. (2023)[40], effectively navigated these challenges by employing pictorial scales and narrative-based stress assessments.

Mixed-age groups were occasionally studied, allowing comparative assessments across developmental stages. The study by Aktas et al. (2022)[28], for instance, covered a broader range of 7–14 years, enabling evaluation across developmental transitions. However, such explicitly comparative studies were infrequent.

Demographic Representation

Overall, demographic diversity regarding neurodivergence, socioeconomic status, and ethnicity was rarely detailed explicitly across studies. A few studies, including Aktas et al. (2022)[28], targeted specific populations such as children with autism spectrum disorder (ASD), demonstrating inclusive research strategies. Nevertheless, broader demographic considerations remain sparse, indicating a critical gap for future research. Notably, a significant proportion of reviewed studies specifically targeted neurodiverse children,

particularly those diagnosed with Autism Spectrum Disorder (ASD). Out of the 49 included studies, 15 (approximately 30.6%) explicitly focused on ASD populations. This strong representation highlights research interest and methodological suitability in applying non-invasive stress-monitoring technologies for neurodiverse groups. However, it also underscores the potential underrepresentation of other neurodivergent populations, as well as typically developing children, within stress-monitoring research.

Table 10. Distribution of studies by environment type (n=49 total)

Environment Type	Number of Studies	Example Citations
Controlled Laboratory Settings	10	Azevedo et al. (2016)[17]; Chiang et al. (2024)[58]; Goulart et al. (2019)[43]; Isbister et al. (2022)[39]; Li et al. (2022)[59]; Melinda et al. (2024)[52]; Nag et al. (2020)[60]; Payani et al. (2021)[55]; Theofanopoulou et al. (2019)[50]; V. P. et al. (2023)[61]
Home or Residential Settings	6	Vural et al. (2024)[37]; Li et al. (2024)[62]; Nguyen et al. (2021)[53]; Redd et al. (2020)[41]; Wang et al. (2025)[56]; Yu et al. (2023)[38]
Not Specified	16	Aktas et al. (2022)[28]; Carlier et al. (2019)[47]; Coşkun et al. (2022)[34]; Coşkun et al. (2023)[29]; Curzio et al. (2022)[63]; Deng et al. (2022)[27]; Gul Airij et al. (2016)[64]; Kumazaki et al. (2017)[44]; Rathod et al. (2022)[65]; Rodriguez-Pellejero et al. (2024)[66]; Sarabadani et al. (2020)[46]; Shonkoff et al. (2022)[67]; Singh (2022)[68]; Thierfelder et al. (2022)[69]; Vitullo and Benitez (2019)[70]; Xiao et al. (2025)[71]
School-Based Settings	17	Lutin et al. (2023)[40]; Antle et al. (2019)[14]; Buddi et al. (2019)[31]; Chaowadee et al. (2021)[32]; Choi et al. (2017)[33]; Chyan et al. (2023)[72]; Coşkun et al. (2023)[35]; Kadirvelu et al. (2023)[30]; Gurusamy et al. (2024)[73]; Zhang et al. (2020)[42]; Jin et al. (2023)[36]; Kim et al. (2021)[48]; Loftness et al. (2022)[45]; Masino et al. (2019)[51]; Nicolaidou et al. (2019)[49]; Northrup et al. (2016)[54]; Yao et al. (2023)[57]

Table 11. Distribution of studies by age group (n=49 total).

Age Group	Number of Studies	Example Citations
Adolescents (13–18 years)	18	Coşkun et al. (2023)[34]; Kadirvelu et al. (2023)[30]; Vural et al. (2024)[37]; Isbister et al. (2022)[39]; Jin et al. (2023)[36]; Kumazaki et al. (2017)[44]; Li et al. (2022)[59]; Loftness et al. (2022)[45]; Melinda

Age Group	Number of Studies	Example Citations
		et al. (2024)[52]; Northrup et al. (2016)[54]; Payani et al. (2021)[55]; Redd et al. (2020)[41]; Rodriguez-Pellejero et al. (2024)[66]; Shonkoff et al. (2022)[67]; Theofanopoulou et al. (2019)[50]; Thierfelder et al. (2022)[69]; Vitullo and Benitez (2019)[70]; Wang et al. (2025)[56]
Children (under 13 years)	18	Lutin et al. (2023)[40]; Antle et al. (2019)[14]; Azevedo et al. (2016) [17]; Carlier et al. (2019)[47]; Chaowadee et al. (2021)[32]; Choi et al. (2017)[33]; Chyan et al. (2023)[72]; Coşkun et al. (2023)[34]; Curzio et al. (2022)[63]; Goulart et al. (2019)[43]; Gul Airij et al. (2016)[64]; Zhang et al. (2020)[42]; Nag et al. (2020)[60]; Nicolaidou et al. (2019) [49]; Rathod et al. (2022)[65]; V. P. et al. (2023)[61]; Xiao et al. (2025) [71]; Yu et al. (2023)[38]
Mixed-Age Pediatric Samples	9	Aktas et al. (2022)[28]; Buddi et al. (2019)[31]; Chiang et al. (2024) [58]; Coşkun et al. (2022)[35]; Gurusamy et al. (2024)[73]; Kim et al. (2021)[48]; Nguyen et al. (2021)[53]; Sarabadani et al. (2020)[46]; Singh (2022)[68]
Unclear or Not Reported	4	Deng et al. (2022)[27]; Li et al. (2024)[62]; Masino et al. (2019)[51]; Yao et al. (2023)[57]

Overall, laboratory and school settings dominated existing studies, likely due to logistical convenience and clear ethical guidelines. In contrast, home-based and hybrid contexts were comparatively rare despite their potential advantages in ecological validity and stress realism. Similarly, adolescent-focused studies significantly outnumbered those involving younger children, leaving developmental stages below 8 years inadequately explored. To address these gaps, future studies should expand the diversity of deployment environments, enhance inclusion of younger and neurodiverse populations, and employ longitudinal designs capable of capturing the dynamic nature of pediatric stress responses in real-world contexts.

3.5. Evaluation Approaches and Performance of Detection Models

The reviewed studies utilized a diverse array of evaluation approaches, reflecting varying degrees of methodological sophistication, interpretability, and practical applicability. We classified the approaches into distinct categories based on the underlying evaluation methodology: traditional statistical methods, supervised machine learning (ML) models, deep learning techniques, personalized or adaptive models, qualitative evaluations, hybrid evaluations, and exploratory or conceptual studies. (Table 12).

Traditional Statistical Methods (Quantitative)

Quantitative methods dominated the evaluation landscape, with 23 out of the reviewed studies employing conventional statistical metrics such as accuracy, precision, recall, F1-score, specificity, sensitivity, and Area Under Curve (AUC). These studies primarily used physiological metrics—Heart Rate Variability (HRV),

Electrodermal Activity (EDA), and accelerometer-derived activity data—to gauge stress levels. Reported accuracies frequently ranged between 82% and 92%, and F1-scores ranged from 79% to 95% (Aktas et al. [28]; Chaowadee et al.[32]; Rodriguez-Pellejero et al.,[66]). Several authors highlighted improved performance through multimodal integration; for instance, Buddi et al. (2019)[31] achieved high predictive accuracy by combining HRV and EDA signals, demonstrating that multimodal data consistently outperforms single-signal approaches.

Supervised Machine Learning Models

Four studies explicitly leveraged supervised ML algorithms, including Support Vector Machines (SVM), Random Forests (RF), and K-Nearest Neighbors (KNN). These techniques showed robust predictive capabilities, with classification accuracies ranging from 86% to 92%, and F1-scores from 81% to 91% (Gurusamy et al.[73]; Zhang et al.[42]). Notably, Xiao et al. (2025)[71] found that multimodal ML approaches outperformed unimodal approaches by leveraging the complementary nature of physiological and contextual features, underscoring the potential for advanced multimodal ML frameworks in pediatric stress detection.

Deep Learning Approaches

Deep learning models, despite their proven performance in adult studies (Arsalan et al., 2023)[13], remain underrepresented in pediatric stress research. Nevertheless, studies such as Chyan et al. (2023)[72] reported compelling results with deep neural architectures—specifically, convolutional (CNN) and recurrent neural networks (RNN)—achieving a remarkable accuracy of 95.6% using vocal stress markers. However, the scalability of these deep learning techniques may be constrained by the high computational resources required and the substantial amount of labeled data necessary to train robust models.

Personalized or Adaptive Models

The adoption of personalized models—those trained specifically on individual physiological and behavioral signatures—remains uncommon, yet the study by Jin et al. (2023)[36] offered significant insights into their potential. Jin and colleagues reported that subject-specific models considerably outperformed generic classifiers, emphasizing the role of personalization in enhancing predictive accuracy and ecological validity, particularly in diverse pediatric populations.

Table 12. Evaluation Methods, Refined Metrics, and Citations

Evaluation Method	No. of Studies	Metrics/Accuracy	Citations
Traditional Statistical Methods (Quantitative)	23	Accuracy: 82–92%, F1-score: 79–95%, precision >80%,	Aktas et al.[28]; Azevedo et al.; Buddi et al.[31]; Chaowadee et al. [32]; Chiang et al.[58]; Choi et al.

Evaluation Method	No. of Studies	Metrics/Accuracy	Citations
		improved HRV	[33]; Chyan et al.[72]; Coşkun et al.[34]; Curzio et al.[63]; Deng et al.[27]; Goulart et al.[43]; Jin et al.[36]; Kim et al.[48]; Kumazaki et al.[44]; Loftness et al.[45]; Masino et al.[16]; Nag et al.[60]; Rathod et al.[65]; Rodriguez-Pellejero et al.[66]; Sarabadani et al.[46]; Thierfelder et al.[69]; V. P. et al. [61]
Supervised Machine Learning Models	4	Accuracy: 86–92%, F1-score: 81–91%, Multi-modal fusion	Gurusamy et al.[73]; Vural et al. [37]; Xiao et al.[71]; Zhang et al. [42]
Deep Learning Approaches	1	Accuracy: up to 95.6% (CNN, RNN-based)	Chyan et al. [72]
Personalized or Adaptive Models	1	Personalized models outperform generic (accuracy improvement)	Jin et al. [36]
Qualitative Evaluation Approaches	12	High usability, improved emotional awareness, subjective satisfaction	Antle et al.[14]; Carlier et al.[47]; Isbister et al.[39]; Kadirvelu et al. [30]; Lutin et al.[40]; Nguyen et al.[53]; Nicolaidou et al.[49]; Northrup et al.[54]; Redd et al. [41]; Vitullo and Benitez[70]; Wang et al.[74]; Yao et al. [57]
Hybrid Evaluations	2	Accuracy: 85–90%, high user satisfaction, feasibility	Theofanopoulou et al.; Yu et al.
Exploratory/Conceptual or Simulation Studies	7	Conceptual demonstration or simulated accuracy (>70%)	Gul Airij et al.[64]; Li et al.; Melinda et al.[52]; Payani et al. [55]; Shonkoff et al.[67]; Singh[68]

3.5.1. Qualitative Evaluation Approaches

Twelve studies utilized qualitative evaluations, incorporating methods such as structured interviews, focus groups, usability questionnaires, and observational analyses. These studies provided valuable insights into

user satisfaction, perceived effectiveness, and practical barriers to implementation. Antle et al. (2019)[14], Kadirvelu et al. (2023)[30], and Lutin et al. (2023)[40], for instance, identified high usability ratings, improved emotional awareness, and substantial acceptance among pediatric users, parents, and caregivers. These findings underscore the critical role qualitative evaluations play in optimizing technological interventions beyond mere numerical performance metrics.

Hybrid Evaluations

Two studies adopted hybrid evaluations, integrating quantitative accuracy assessments with qualitative feedback from users (Theofanopoulou et al.[50], 2019; Yu et al., 2023[38]). These multidimensional evaluations captured both statistical accuracy (85%–90%) and high user satisfaction, thus providing a comprehensive view of system performance and acceptability within real-world contexts. Such integrative methods offer a powerful framework for future evaluations aiming for both technical rigor and user-centered validation.

Exploratory, Conceptual, or Simulation-Based Studies

Seven studies (Melinda et al., 2024[52]; Payani et al.[55], 2021; Singh, 2022[68]; Li et al., 2024[62]; Gul Airij et al.[64], 2016; Shonkoff et al., 2022[67]) explored initial feasibility, conceptual frameworks, or early-stage simulations without explicit validation against physiological ground truth. These studies typically demonstrated conceptual viability, preliminary effectiveness, or simulated accuracy greater than 70%, though lacking the rigor of controlled evaluation scenarios. Nevertheless, these exploratory approaches provide foundational insights for future systematic validation efforts.

This comprehensive evaluation landscape demonstrates clear trade-offs between methodological rigor, practical interpretability, and ecological validity. Traditional statistical and supervised ML approaches currently dominate pediatric stress detection research, consistently achieving moderate to high accuracy. However, personalized and deep learning methods, although less explored, show substantial promise in enhancing predictive performance and adapting to individual differences. Qualitative and hybrid evaluations further enrich our understanding by emphasizing user-centered outcomes critical for real-world adoption.

Future research should pursue more personalized adaptive modeling, standardized evaluation protocols, expanded deep learning exploration, and increased use of hybrid methodologies to bridge quantitative robustness with qualitative user-centered validation. Longitudinal studies and broader participant representation will also be critical in confirming the generalizability and practical utility of pediatric stress-monitoring technologies. Notably, the highest reported accuracy (95.6%) was achieved using deep learning on vocal stress markers in a controlled laboratory setting [72], whereas most statistical and supervised ML approaches based on HRV and EDA achieved accuracies in the 82–92% range. This highlights both the potential and the limitations of current evaluation approaches.

3.6. Stress Datasets and Data Collection Methodologies

High-quality datasets are foundational for the development and validation of stress-monitoring technologies. They enable comparative benchmarking, training of predictive models, and replicability of findings. In the context of pediatric stress detection, datasets are especially critical due to the nuanced nature of children's behavioral and physiological responses, as well as the ethical complexities involved in data collection. This section analyzes the types of datasets used across the reviewed studies, their signal modalities, annotation methods, and availability.

Types and Availability of Datasets

Out of the 49 studies reviewed, 48 involved some form of dataset collection, though only 7 explicitly reported using publicly available datasets. These public datasets were often labeled vaguely—e.g., “*Custom (shared)*”, “*Own*”, or “*From their 2023 database*”—without references to standardized repositories such as DEAP or WESAD. This suggests that while public access was mentioned, accessibility and reusability remain limited. Despite a few instances of data sharing, the field still lacks standard, openly accessible pediatric stress datasets, which hinders reproducibility and large-scale model training.

Data Modalities Captured

The datasets reviewed varied widely in the physiological, behavioral, and contextual signals recorded. *Data modality* refers to the type of input signals collected to infer stress, including physiological signals such as Heart Rate Variability (HRV), Electrodermal Activity (EDA), respiration, temperature, and EEG; behavioral signals such as facial expressions, posture, and voice tone; and contextual data including location, device usage patterns, or task metadata.

Physiological signals were the most common modality, featured in over 70% of data-collecting studies (e.g., Buddi et al., 2019[31]; Chaowadee et al., (2021)[32], often using wearable sensors like Empatica E4 or Garmin HR monitors. Behavioral signals were integrated in studies such as Antle et al. (2019)[14], which coupled EEG and observational feedback in a school-based biofeedback system. Few studies incorporated contextual data, although examples exist where task events or school routines served as time-based proxies (e.g., Zhang et al., 2020[42]).

Labeling and Annotation Approaches

Labeling, or annotation, refers to the method by which ground truth stress states were established for model training or analysis. Annotation quality significantly influences model validity, especially when working with populations like children who may not reliably self-report internal states.

Three main labeling strategies were observed:

1. **Self-report-based labeling:** Common in older children and adolescents, used Likert scales, stress thermometers, or emotional state selection tools (e.g., Lutin et al.[40], 2023; Azevedo et al., 2016[17]).

2. **Expert annotation:** Used in clinical and therapy studies where behavioral coders or therapists labeled sessions based on predefined stress criteria (e.g., Aktas et al., 2022[28]; Vural et al., 2024).
3. **Event-based labeling:** Stress labels were inferred from task timing, such as exam simulations, robot interaction sessions, or public speaking challenges (e.g., Zhang et al., 2020[42]; Chyan et al., 2023[72]).

Population and Sample Characteristics

Sample sizes ranged from 8 to 84 participants, with a median close to 30. Several datasets focused on children with Autism Spectrum Disorder (ASD) or other neurodevelopmental conditions (e.g., Carlier et al., 2019[47]; Aktas et al., 2022[28]), while others included typically developing children in school environments. Home settings were notably underrepresented in data collection despite their relevance for ecological validity.

Duration and Context of Data Collection

Most datasets were collected during single-session lab studies or short-term school deployments. Only a handful (e.g., Zhang et al., 2020[42]) collected longitudinal data over multiple weeks. Lab-based data tended to prioritize physiological precision, while school- and therapy-based data offered greater ecological realism.

Commonly collected modalities included physiological signals such as HRV, EDA, EEG, and accelerometry, while fewer datasets incorporated audio or contextual data. Annotation methods predominantly relied on self-report, especially in school-aged children, followed by expert tagging in clinical contexts and event-based labels in task-driven studies. The most frequent data collection environments were school-based, clinical, or lab-controlled settings. While some studies reported publicly available datasets, most were proprietary or lacked clear accessibility. The field currently lacks standard, openly accessible pediatric stress datasets with consistent annotation and broad demographic coverage. Addressing these gaps is essential for enabling reproducibility, longitudinal modeling, and broader generalizability in stress-monitoring research.

3.7. Mitigation Strategies and Interventions

While many stress-monitoring systems aim to detect or classify stress in children, fewer go beyond detection to include strategies for real-time stress reduction, emotional regulation, or behavioral adaptation. This section synthesizes how the reviewed studies incorporated mitigation strategies—whether through biofeedback, adaptive interfaces, or supportive therapeutic tools—and evaluates the underlying technologies, response modalities, and reported impacts. (Table 13).

Table 13. Summary of mitigation strategies in pediatric stress studies (n=9). Most interventions were short-term in duration (single session to a few weeks) and were deployed in settings such as schools, homes, or therapeutic centers.

Study	Mitigation Type	Signals Used	Real-Time or Offline	Reported Impact
Antle et al. (2019)[14]	EEG-based biofeedback	EEG (alpha)	Real-time	Improved self-regulation
Carlier et al. (2019) [47]	Serious game with adaptive stress response	HR, EDA	Offline (session)	Reduced arousal, improved engagement
Theofanopoulou et al. (2019)[50]	Wearable HRV feedback	HRV	Real-time	Improved awareness, relaxation
Wang et al. (2025)[56]	Smartwatch with alert feedback	HR	Real-time	Stress awareness improvement
Yu et al. (2023)[38]	Adaptive mobile interface	HRV, interaction context	Real-time	Enhanced self-awareness, behavior change
Chyan et al. (2023[72])	Vocal stress-adaptive system	Audio (RNN)	Real-time	Improved user engagement
Kadirvelu et al. (2023) [30]	Parent-facing support interface	HR, behavioral cues	Near real-time	Improved caregiver guidance
Redd et al. (2020)[41]	Wearable relaxation prompt	EDA, HR	Real-time	Reduced anxiety, better coping
Lutin et al. (2023)[40]	School HR tracking dashboard	HR	Offline	Preventive stress insight in classrooms

Types of Mitigation Strategies

Among the 49 reviewed studies, 9 explicitly implemented mitigation or intervention features aimed at reducing stress or enhancing emotional self-regulation. These approaches fell into three major categories:

1. **Biofeedback systems:** Tools that provide real-time visual, auditory, or haptic cues based on physiological signals to support relaxation or awareness. Antle et al. (2019)[14], for instance, employed EEG-based feedback to support emotional self-regulation in school environments. Similarly, Theofanopoulou et al. (2019)[50] and Wang et al. (2025)[56] used wrist-worn devices that provided visual or tactile feedback based on real-time HRV changes.
2. **Game-based or interaction-centered interventions:** Carlier et al. (2019)[47] demonstrated that serious games with embedded stress-adaptive components helped reduce social anxiety and stress among children with Autism Spectrum Disorder (ASD). Chyan et al. (2023)[72] explored a voice-sensitive RNN-based system that adapted content delivery in response to real-time vocal stress indicators.
3. **Context-aware adaptive interventions:** Yu et al. (2023)[38] introduced a mobile-based adaptive application that dynamically adjusted its interface based on physiological and contextual stress patterns. Kadirvelu et al. (2023)[30] and Redd et al. (2020)[41] emphasized caregiver-oriented support tools that delivered behavior cues and recommended calming actions in response to detected stress episodes.

Signals and Modalities Used

Most mitigation systems utilized physiological signals such as HR, HRV, and EEG. Biofeedback loops were most often powered by HRV changes (Wang et al., 2025[56]; Theofanopoulou et al.[50], 2019) or EEG alpha band shifts (Antle et al., 2019[14]). Behavioral signals—such as vocal prosody, posture, or facial cues—were less commonly used in feedback design but featured in exploratory studies (e.g., Chyan et al., 2023[72]).

Real-Time vs Offline Adaptation

Roughly half of the intervention studies employed real-time mitigation (e.g., smartwatches, mobile apps), while others relied on session-based post hoc modulation (e.g., game-based design with feedback administered after detection). Real-time systems had greater potential for behavioral reinforcement but required more sophisticated processing and battery-aware design.

Reported Impacts

Most mitigation studies reported improvements in emotional regulation, self-awareness, or reduced anxiety markers. For example, Carlier et al. (2019)[47] observed reduced physiological arousal and improved task engagement following game-based sessions. Yu et al. (2023)[38] and Redd et al. (2020)[41] noted increased parental awareness and reported behavior improvements in longitudinal contexts.

Despite promising directions, mitigation-focused systems remain underexplored in pediatric stress technology. Biofeedback remains the dominant strategy, but long-term validation, age-personalized interfaces, and behaviorally anchored feedback are still limited. Future work should explore adaptive learning, multi-user interventions (e.g., child + caregiver), and scalable delivery platforms such as wearables or smart environments.

4. Discussion

This scoping review provides a comprehensive synthesis of non-invasive technologies developed for stress detection and management in pediatric populations. Anchored in seven structured research questions, the review reveals not only the field's notable technological and methodological advancements but also several persistent limitations that hinder real-world adoption and generalizability.

Key challenges span four interconnected domains: population representation, cultural and contextual adaptation, technical and methodological consistency, and ethical and privacy safeguards. Together, these limitations critically affect the ecological validity, inclusiveness, and practical utility of current systems. Addressing these gaps is essential to ensure that stress-monitoring technologies are not only scientifically robust but also ethically grounded, culturally relevant, and adaptable to the diverse realities of children's lives across global contexts.

4.1. Non-Contact Methods in Pediatric Stress Detection

The field of pediatric stress detection has seen a substantial increase in the development of non-invasive

technologies. However, the vast majority of these systems continue to rely on contact-based sensing modalities such as wrist-worn wearables, chest straps, and EEG headsets. In our review of 49 studies, 41 (approximately 84%) employed contact-based systems, 7 (14%) used clearly non-contact methods, and only one study utilized a hybrid approach combining both.

Contact-based devices, particularly those capturing HRV, EDA, or temperature data, remain popular due to their compatibility with established physiological stress markers and existing signal processing pipelines. Wearables such as Empatica E4, BioPlux, and Fitbit devices were frequently deployed [28,29,31,33]. However, their design often fails to accommodate pediatric-specific needs. Carlier et al. [47] and Antle et al. [14] reported compliance issues stemming from the physical burden of wearables, especially among children with ASD or sensory processing sensitivities.

By contrast, non-contact modalities—though underrepresented—demonstrated growing versatility. Of the seven studies employing non-contact approaches, four used vision-based systems. Goulart et al. [43] implemented thermal and RGB facial analysis during robot-child interaction; Kumazaki et al. [44] used camera-equipped humanoid robots to monitor behavioral stress responses in children with ASD; Coşkun et al. [35] applied video observation with clinical annotation in serious game sessions; and Yu et al. [38] utilized ambient mobile logging coupled with caregiver-tagged events. Additionally, one study used audio-based vocal analysis [72], and two others included environmental sensors but lacked biosignal specificity.

Performance comparisons revealed that contact-based systems—especially those using multimodal physiological sensing—achieved classification accuracies above 90% [27,72,75]. Reported accuracies for non-contact systems (68–88%) varied with lighting, occlusion, and camera angle constraints in controlled settings; none of the non-contact studies conducted long-term, in-the-wild deployments. These constraints—such as lighting variability, camera angle, and occlusion—were identified as major performance bottlenecks [34,43].

While several non-contact studies provided strong conceptual validation, all were conducted in controlled environments. None implemented long-term or real-world monitoring in classrooms or homes. This represents a critical limitation, as stress behaviors in children may differ substantially across settings.

Notably, facial action coding systems (e.g., OpenFace), gaze tracking, or pose estimation were rarely used—despite their potential to offer non-invasive insights into emotional states. Furthermore, contactless biosignal acquisition methods such as remote photoplethysmography (rPPG) or microwave reflectometry, though effective in adult stress studies [16], were not found in any of the reviewed pediatric works.

Hybrid sensing remains underutilized. Only one study [72] combined non-contact (voice) and contact (movement sensor) features in a unified classification model. Given the complementary strengths of each modality, this gap highlights missed opportunities to improve robustness and usability.

To advance non-contact sensing in pediatric stress research, future studies should:

- Compare contact and non-contact systems within the same populations and tasks.
- Implement vision- or audio-based monitoring in naturalistic school and home settings.
- Apply data fusion techniques combining facial, vocal, and contextual signals.
- Design privacy-aware, culturally adaptive systems that account for developmental ergonomics.

Overall, while contact-based systems currently dominate the field, non-contact approaches—when carefully designed—can offer meaningful advantages in terms of comfort, compliance, and real-world adaptability. Future work must prioritize this underexplored space to enable inclusive, scalable pediatric stress monitoring.

4.2. Stress Induction and Validation Protocols

The reviewed studies revealed a diverse array of stress induction paradigms used to elicit measurable responses in children. These included cognitive challenges (e.g., memory tasks, timed quizzes), social-evaluative stressors (e.g., robot-mediated interviews, public speaking simulations), game-based challenges (e.g., escalating difficulty levels), and passive observation in naturalistic environments (e.g., daily classroom routines or home settings). This diversity reflects efforts to mirror stressors encountered in daily life, but also underscores the absence of a standardized framework for elicitation and validation.

A notable gap lies in the limited use of standardized protocols established in broader stress research. The Trier Social Stress Test (TSST), considered a gold standard in adult studies for inducing psychosocial stress, was not adapted in any of the reviewed pediatric works. Similarly, validated affective stimuli—such as the International Affective Picture System (IAPS) or emotion-inducing video clips—were entirely absent from the corpus. The lack of these reproducible and benchmarked stressors hinders comparability across studies.

Cognitive and academic stressors were most common across the 49 studies. Buddi et al. [31], Zhang et al. [42], and Jin et al. [36] used exam-like tasks and time-pressured quizzes. Social-evaluative stressors, although fewer, included robot-mediated conversation [44], mock interviews [45], and scripted interactions [43]. Game-based paradigms introduced stress through escalating task difficulty or competition, as seen in Carlier et al. [47], Coşkun et al. [34], and Nicolaidou et al. [49]. (Table 6).

Approximately one-quarter of the studies (e.g., [38,40,41]) adopted naturalistic stress monitoring, eschewing explicit induction in favor of passive observation or caregiver-annotated episodes. While this improves ecological validity, it poses clear challenges for reproducibility and cross-study benchmarking.

Validation methods were equally heterogeneous. Self-report tools—ranging from stress thermometers to Likert scales and pictorial SAM instruments—were widely used, though their reliability in young or neurodiverse populations remains limited [40]. Physiological indicators like HRV and EDA were consistently used in contact-based systems [14,27,28]. Fewer studies employed observer-based behavioral coding or synchronized video analysis [34,43].

However, methodological inconsistencies in annotation, stress windowing, and validation metrics were common. Most studies did not report effect sizes, statistical power, or inter-rater reliability. Moreover, stress

induction intensities and validation sensitivity were rarely evaluated across demographic or developmental groups.

To strengthen future work, we recommend the following:

- Adapt and validate standardized protocols like the TSST for use in children.
- Combine physiological and behavioral validation in multimodal frameworks.
- Establish age-appropriate, standardized annotation schemes.
- Report repeatability, effect sizes, and context-specific sensitivity.

A more rigorous and harmonized approach to stress induction and validation will improve interpretability, reproducibility, and comparability across pediatric stress monitoring studies.

4.3. Contextual, Demographic and Cultural Considerations

The deployment contexts and target demographics of pediatric stress-monitoring systems emerged as significant limitations in the reviewed literature. A substantial number of studies were conducted in controlled laboratory or school settings. While laboratories offer controlled environments ideal for validating sensor accuracy and reproducing experimental procedures, they do not adequately reflect the complexity and unpredictability of real-world stressors, especially for younger children whose behaviors may vary drastically outside such environments (Lutin et al., 2023[40]; Antle et al., 2019[14]). School-based studies, although more ecologically valid, often suffered from short observation periods and lacked longitudinal designs necessary for capturing the dynamic nature of pediatric stress responses (Buddi et al., 2019[31]; Zhang et al., 2020[42]).

Demographically, the literature exhibited a strong skew toward adolescents, with children under 8 years significantly underrepresented. This imbalance is likely due to practical challenges, such as lower compliance with wearables and limited verbal articulation of internal states in younger children (Carlier et al., 2019[47]; Redd et al., 2020[41]). However, one promising trend was the relatively strong representation of children with Autism Spectrum Disorder (ASD). At least five studies—such as those by Aktas et al. (2022)[28], Carlier et al. (2019)[47], Coşkun et al. (2023)[34], Kumazaki et al. (2017)[44], and Yu et al. (2023)[38]—specifically targeted children with ASD, exploring both stress detection and intervention strategies tailored to their unique behavioral and sensory needs. This emphasis is encouraging and suggests that researchers are recognizing the methodological importance and clinical value of working with neurodiverse populations.

Furthermore, a critical limitation was the lack of broader demographic diversity. Many studies used small, homogeneous samples, often drawn from urban, middle-to-high socioeconomic backgrounds, limiting the external validity and generalizability of findings (Jin et al., 2023[36]; Deng et al., 2022[27]). For example, Buddi et al. (2019) noted recruitment challenges in their rural Indian deployment, underscoring the need for locally adapted solutions and inclusive recruitment strategies. This contrasts starkly with adult stress detection literature, which more frequently engages broader and more demographically varied populations (González Ramírez et al., 2023[12]; Aqajari et al., 2023[20]).

Closely linked to these demographic challenges are cultural and contextual blind spots. While some studies acknowledged the importance of contextual fit—such as localized deployments or adaptations for children from lower socioeconomic backgrounds—in-depth cultural tailoring was generally lacking. Studies employing robots, gamified platforms, or serious games frequently adopted culturally agnostic designs that may not align with the behavioral norms, communication patterns, or stress triggers relevant to children from diverse backgrounds. Buddi et al. (2019)[31], for instance, implemented DetectStress; a contextual customization for an Indian rural school setting, addressing linguistic and contextual nuances specific to their deployment. A few other studies similarly acknowledged cultural elements in their methodology. However, most overlooked comprehensive cultural adaptation, thereby risking reduced efficacy, engagement, and acceptance when technologies are implemented across diverse populations or geographical contexts. Thus, rigorous cross-cultural validation and customization remain significantly underexplored areas demanding focused research.

To move toward scalable and equitable applications, future research must adopt deliberate efforts to improve inclusivity. This includes recruiting from underserved or underrepresented populations, designing culturally and developmentally appropriate technologies, and conducting extended field trials in real-world environments like homes and community centers. Cultural adaptation should be integrated into both the design and validation processes to enhance engagement, relevance, and long-term adoption. Such steps are vital to ensure that non-invasive stress-monitoring systems can be generalized across the full spectrum of pediatric users and contexts.

4.4. Methodological and Technical Rigor in Stress Detection

A wide spectrum of technological and methodological approaches has been adopted to detect stress in children, yet considerable challenges persist in ensuring their practical, scalable, and ethically robust deployment.

4.4.1. Technical Constraints and Model Scalability

Many studies leveraged machine learning (ML) classifiers such as Support Vector Machines (SVM), Random Forests (RF), and K-Nearest Neighbors (KNN), often applied to physiological data like heart rate variability (HRV) and electrodermal activity (EDA). These models reported accuracies ranging from **82% to 92%** and F1-scores from **79% to 95%**, especially when multimodal inputs were integrated (e.g., Buddi et al.[31], 2019; Choi et al., 2017[33]; Gurusamy et al., 2024[73]).

However, the adoption of more sophisticated deep learning techniques—such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs)—while showing *accuracy up to 95.6%* using vocal stress markers (Chyan et al., 2023[72]), remained rare due to their computational complexity and need for large labeled datasets. These demands make real-time deployment difficult in school or home environments where device constraints are common.

Personalized models, though promising, were even less explored. Jin et al. (2023)[36] demonstrated that *subject-specific models significantly outperformed generic classifiers*, particularly in detecting nuanced stress responses in adolescents. Yet, their reliance on individualized calibration and longitudinal data collection limits feasibility at scale.

4.4.2. Methodological Inconsistencies and Evaluation Gaps

Beyond computational demands, the field suffers from a lack of methodological standardization. Studies varied in their labeling strategies: some relied on self-reports or pictorial scales (Lutin et al., 2023)[40], while others used *expert-coded behavior* (Aktas et al., 2022[28]; Vural et al., 2024[37]) or event-based proxies from structured tasks like exams and robot interactions (Zhang et al., 2020[42]; Chyan et al., 2023[72]).

This heterogeneity undermines reproducibility and limits comparability across studies. Many failed to report cross-validation procedures or did not benchmark against public datasets. Although some datasets were shared, most remained proprietary, and very few were pediatric-specific, further stalling progress in generalizable model training.

To close these gaps, the field must prioritize:

- Development of *benchmark datasets* with consistent pediatric annotations.
- Use of *standardized performance metrics* (AUC, F1-score, sensitivity/specificity).
- Clear documentation of *validation strategies*, including holdout sets or cross-context evaluations.
- Exploration of *lightweight, interpretable models* that can run on mobile or edge devices.

With these improvements, stress detection systems can move beyond experimental prototypes toward deployment-ready tools for real-time monitoring and intervention in educational and therapeutic contexts.

4.5. Dataset Challenges and Opportunities

Datasets are the cornerstone of reliable stress detection technologies. Several studies have demonstrated innovation in data collection protocols by incorporating physiological, behavioral, and contextual signals. Notably, recent datasets such as StressID [76] and MMSD [77] represent a shift toward richer, multimodal pediatric datasets that include biosignals (e.g., ECG, EDA), audio/video inputs, and caregiver-labeled stress events. Some studies have even begun to integrate salivary cortisol or subjective stress scales alongside sensor data, enriching the multidimensional understanding of stress.

Encouragingly, annotation strategies are evolving to be more developmentally appropriate. For example, Lutin et al. [40] used pictorial stress thermometers and child-friendly narrative prompts to assist in stress reporting. Azevedo et al. [17] applied behavioral tagging derived from posture and task performance to infer stress. These studies illustrate the field's increasing awareness of the cognitive and communicative limitations of younger children, and the need to move beyond adult-centric labeling tools.

Despite these advances, critical limitations remain. Most studies reviewed relied on small sample sizes—often fewer than 40 participants—limiting statistical validity and the diversity of stress profiles represented.

Publicly available pediatric stress datasets are still rare. Datasets like WESAD [75] and StudentLife [78], while influential, were designed for adults and lack child-specific stress indicators or labeling frameworks. Even valuable repositories such as the Pediatric Heart Network focus mainly on physiological baselines and omit contextual or emotional annotations.

Annotation inconsistency also persists. Labels are frequently derived from self-reports, expert video coding, or event-based task windows—each with trade-offs. Self-reports may be unreliable for younger or neurodiverse children. Expert coding, while rigorous, is time-intensive and costly. Event-based annotation (e.g., exam period, game level) assumes a uniform stress response, which may not hold across diverse children or contexts.

Ecological limitations further restrict model development. Most datasets are gathered in constrained environments like labs or classrooms, with limited duration and natural variability. Longitudinal, real-world datasets remain scarce, and models trained in one context often fail to generalize to others. For example, Prajod et al. [79] found that HRV-based models trained on WESAD performed poorly in everyday stress scenarios unless context-specific harmonization was applied.

Nonetheless, the field is steadily progressing. Some emerging efforts have adopted open-data sharing with anonymization, promoted diversity in recruitment, and incorporated ethical design from the outset. Research on domain adaptation (e.g., ComBat harmonization) is beginning to surface as a promising avenue to improve cross-dataset generalization [80].

Looking ahead, it is essential to invest in the development of pediatric-specific, context-rich, and ethically shared datasets. These should be annotated using standardized, multimodal protocols and collected across real-world environments and developmental stages. Doing so will catalyze the creation of robust, inclusive, and scalable stress detection systems for children.

Encouraging open access to systematically annotated, diverse, and longitudinal pediatric datasets will profoundly advance methodological robustness and promote collaborative validation efforts across the scientific community.

4.6. Intervention and Mitigation Strategies

Stress mitigation is a crucial component of any pediatric monitoring system. While detection enables recognition, interventions are necessary to support children in regulating their emotional and physiological responses. The reviewed literature outlines various strategies, including biofeedback, serious games, mobile applications, and caregiver-oriented tools. While these demonstrate promise, systematic evaluation and longitudinal evidence remain limited.

Biofeedback was a core feature in multiple interventions. Antle et al. [14] integrated real-time physiological feedback via an anxiety-focused mobile app, allowing children to monitor their heart rate and electrodermal

activity during gameplay. The system aimed to improve self-awareness and was favorably received in short-term school deployments. Similarly, Theofanopoulou et al. [50] deployed a wrist-worn device that delivered HRV-based cues in classroom settings, supporting attentional regulation through direct physiological feedback.

Game-based interventions were also common. Carlier et al. [47] designed a serious game tailored for children with Autism Spectrum Disorder (ASD), embedding adaptive stress challenges. The immersive environment promoted engagement while targeting emotion regulation. Coşkun et al. [34] developed a therapy-integrated game that used physiological monitoring (e.g., HR, EDA) to classify and respond to stress. These examples support the potential of gamified strategies to align therapeutic objectives with child-centric interaction.

Context-aware and caregiver-facing mobile platforms are emerging as scalable solutions. Yu et al. [38] introduced a mobile app that adapted its interaction design based on physiological and contextual signals. Kadirvelu et al. [30] created a parent-facing system that provided stress insights and recommended coping responses based on passive behavioral sensing. These designs reflect a shift toward caregiver-involved, multi-user intervention models.

Despite this progress, notable gaps remain. Most systems were tested in short-term trials and rarely evaluated in naturalistic or longitudinal deployments. Scalability was also a concern, as many required clinical-grade sensors or manual setup—conditions that may not be feasible in real-world, resource-constrained environments. In addition, usability metrics—especially from the perspective of neurodiverse or underserved children—were inconsistently reported, limiting conclusions on accessibility and long-term adherence.

Future research should address these limitations through:

- Long-term studies across varied environments (e.g., homes, schools, clinics).
- Emphasis on co-design and usability testing with both children and caregivers.
- Prioritization of mobile-ready, low-cost solutions suited to everyday pediatric care.
- Evaluation frameworks that capture both behavioral and physiological outcomes.

In conclusion, intervention strategies in pediatric stress monitoring are growing in sophistication, but their translational potential depends on deeper, context-aware validation and sustained usability in real-world settings.

4.7. Ethical, Privacy, and Scalability Challenges in Real-World Implementation

While the advancement of pediatric stress detection technologies has been notable, the ethical, privacy, and scalability dimensions remain underexplored and unevenly addressed. Across the literature, studies consistently mention ethical considerations such as informed consent, assent, data sensitivity, and privacy. However, these discussions are often surface-level, with few studies implementing or evaluating rigorous frameworks for data governance or child-centered communication.

Physiological and behavioral data—including heart rate variability, galvanic skin response, facial expressions,

and vocal prosody—are inherently sensitive. Without strong data minimization, encryption, or anonymization practices, there is a risk of stigmatization or misuse—particularly in school or clinical environments. Furthermore, younger children may not fully understand what is being measured or why. As highlighted by the U.S. National Institutes of Health [6], ethical pediatric research must ensure comprehension and minimize psychological impact, especially when biosignals are collected continuously.

Several authors have called for stronger ethical oversight. Nicolaidou et al. [49] emphasized the need for child-friendly design and transparency in mobile mental health apps, while Menéndez and Bello Orgaz [81] underscored the importance of formal frameworks for consent and data sharing. However, very few studies implemented third-party ethical audits, age-adapted consent materials, or real-time feedback to participants—practices increasingly common in digital mental health research [44].

Another critical barrier is scalability. Although systems based on wearables, computer vision, or biofeedback show high accuracy in controlled environments, many require costly sensors, cloud processing infrastructure, or clinician supervision. These resource demands limit deployment in rural, low-income, or overcrowded public school settings. For example, Chyan et al. [72] demonstrated impressive deep learning performance using audio signals, but the model's real-time inference required GPU support and high-bandwidth conditions.

Few studies evaluated interventions in naturalistic settings over extended periods. Most lacked cost analyses or sustainability assessments. As MedTech News [15] notes, wearable health technologies for children often fail to transition beyond pilot studies due to manufacturing costs, maintenance, and caregiver resistance. Based on our synthesis, non-invasive pediatric stress-monitoring systems should incorporate:

- **Data minimization and on-device inference:** collect only essential signals; process locally where possible to reduce privacy risk.
- **Transparent and child-appropriate consent/assent:** use age-adjusted explanations, parental opt-in, and ongoing assent checks.
- **Role-based access and retention policies:** limit data visibility to caregivers/clinicians with clear retention and deletion schedules.
- **Privacy-by-design defaults:** no raw storage of audio/video; anonymization and encryption for physiological/behavioral signals.
- **Equity and accessibility:** ensure devices and apps are affordable, culturally adaptable, and usable in low-resource school and home contexts.
- **Cost-effectiveness and sustainability:** prioritize low-cost, modular hardware and integration into existing educational and telehealth infrastructures.

5. Implications for Research and Practice

5.1. Implications for Researchers

Standardize pediatric elicitation and validation. The review shows wide variability in stressors (cognitive, social/robot-mediated, game-based, naturalistic) and labeling (self-report, expert coding, event-based), with no pediatric adaptation of canonical protocols like TSST. Establish age-appropriate protocols, multi-modal validation (physiological + behavioral), and reporting standards (AUC, F1, sensitivity/specificity, power/effect

sizes).

Prioritize ecological validity. Most deployments occur in schools and labs; home and hybrid in-the-wild studies are comparatively rare and short-term. Future work should emphasize longitudinal, cross-context datasets (home–school–clinic) to capture day-to-day variability and reduce lab bias.

Broaden age and population coverage. Adolescents dominate samples; children below 8 years and underserved populations are underrepresented, while ASD is relatively well covered. Methods, instruments, and assent procedures should be tailored for younger cohorts and for neurodiversity beyond ASD, with explicit recruitment for demographic diversity.

Invest in open, pediatric-specific benchmarks. Public pediatric stress datasets remain scarce; annotation schemes and accessibility are inconsistent. Community benchmarks with shared protocols (signals, labels, contexts) are needed to enable reproducible comparisons and model generalization.

Modeling direction: pragmatic + personalized. Typical statistical and supervised ML approaches yield 82–92% accuracy (F1 79–95%), but datasets and cross-study generalization are limits; deep learning shows promise (e.g., high accuracy on vocal markers) yet is data- and compute-intensive. Favor lightweight models that run on-device; evaluate personalization where feasible; and report cross-dataset transfer.

Shift toward non-contact and hybrid sensing. Contact-based systems dominate (41/49; ~84%); non-contact systems are underused and tested mainly in controlled settings with variable accuracy (68–88%). Head-to-head comparisons and fusion of facial, vocal, contextual, and minimal physiological signals can improve comfort and compliance for children.

Make mitigation first-class. Only 9 of 49 studies implemented interventions (biofeedback, adaptive interfaces, caregiver tools), often short-term. Future studies should embed closed-loop mitigation, measure child-centered outcomes (awareness, coping), and evaluate durability in real-world use.

Operationalize ethics and scalability. Move beyond statements of consent/assent to privacy-by-design (data minimization, on-device processing), child-appropriate disclosures, and cost/maintenance analyses for resource-constrained settings.

5.2. Implications for Industry and Practice

Schools and ed-tech. Near-term value lies in classroom-friendly tools that provide low-friction monitoring and feedback (e.g., HR/HRV dashboards), with configurable thresholds, teacher/caregiver views, and no raw data export by default. Pilot evidence suggests dashboards and prompts can raise awareness and support coping.

Healthcare and telehealth. Clinically aligned systems should integrate interpretable indicators (HRV, EDA) and caregiver-facing guidance, emphasizing alert fatigue reduction and integration with existing workflows rather than stand-alone apps.

Devices and platforms. Favor minimal sensors and edge inference to reduce cost, friction, and privacy exposure. Where cameras or microphones are used, incorporate privacy-safe non-contact pipelines (no storage of raw audio/video; features only) and fallbacks for lighting/occlusion.

Deployment governance. Provide implementation kits: role-based access, retention schedules, child-friendly consent artifacts, incident playbooks, and calibration guides for different age groups and sensory profiles.

For Clinicians & Implementers. To support adoption of non-invasive pediatric stress-monitoring tools in practice, implementations should emphasize:

- **Integration into workflow:** align with existing school routines and telehealth platforms rather than adding standalone apps.
- **Reducing alert fatigue:** provide tiered alerts or summary dashboards to caregivers/teachers instead of constant notifications.
- **Bring-your-own-device (BYOD) constraints:** design mobile apps and wearables that function across heterogeneous devices common in schools and homes.
- **Privacy-preserving defaults:** ensure no raw video/audio storage, clear opt-ins, and caregiver visibility only where appropriate.
- **Actionable feedback:** deliver stress insights in child- and caregiver-friendly formats (visual cues, traffic-light dashboards, or simple reports).

5.3. Research Directions Opened by This Review

1. **Pediatric protocol suite.** Develop and validate a reproducible suite of age-graded stressors and annotation tools (pictorial/self-report + expert coding), with required metrics and reporting checklists.
2. **Longitudinal, cross-context cohorts.** Build multi-week home-school datasets with standardized labels and context logs to study stability, transfer, and seasonality.
3. **Non-contact + hybrid pipelines.** Benchmark rPPG/vision/audio against wearables in the same tasks and fuse modalities for robustness under real-world constraints.
4. **Young-child methods.** Design ergonomics, assent, and feedback for 18 years; evaluate pictorial and narrative tools and reduce sensor burden.
5. **From detection to support.** Expand closed-loop interventions (child + caregiver) with real-time prompts and measure outcomes beyond accuracy (coping skills, school participation, caregiver confidence) over months.
6. **Ethics, privacy, cost.** Pre-register privacy budgets, on-device processing plans, and cost models; include third-party audits and equity reporting for underserved settings.

6. Threats to Validity

While this scoping review provides a comprehensive synthesis of non-invasive technologies for stress detection and management in children, several limitations should be considered when interpreting the findings.

First, although a structured search strategy was implemented, the review was restricted to English-language peer-reviewed literature. This may have led to the exclusion of relevant studies published in other languages or within grey literature domains (e.g., theses, dissertations, industry reports), thereby introducing potential selection bias.

Second, the classification of technologies—particularly when distinguishing between contact-based, non-contact, and hybrid systems—involved subjective judgment based on available methodological descriptions. While predefined coding rules were applied, inconsistencies in how authors described sensor placement, usage context, or intended modality may have influenced categorization. Additionally, the assignment of themes such as “intervention” or “validation strategy” relied on nuanced interpretation, introducing a risk of interpretive bias.

Third, reporting quality varied across studies. Several papers lacked detailed descriptions of sample demographics, stress induction protocols, or evaluation methods, limiting the ability to assess comparability. In some cases, metrics such as accuracy, sensitivity, or F1-score were either not reported or inconsistently defined, complicating synthesis and cross-study benchmarking.

Fourth, while the review aimed for inclusivity, its methodological scope—focusing on technologies explicitly labeled as “non-invasive”—may have inadvertently excluded adjacent domains such as mobile health platforms or passive behavioral sensing tools not categorized under stress detection. Furthermore, the inclusion cutoff date (2025) means that emerging post-review studies are not reflected.

Lastly, the heterogeneity of the field presents both a strength and a challenge. Variability in age groups, use settings (e.g., school vs. clinical), modalities (e.g., EDA vs. thermal imaging), and outcomes of interest (e.g., stress classification vs. behavior tracking) limits the generalizability of specific conclusions and underscores the need for standardization across future studies.

Despite these limitations, we mitigated validity threats through systematic screening, consistent coding procedures, and the inclusion of diverse study types and settings. The review offers a grounded and transparent account of the current landscape and serves as a reference point for future work aiming to design inclusive, effective, and context-sensitive pediatric stress monitoring systems.

7. Conclusion

Amid growing concerns about children’s mental health and emotional resilience, this scoping review offers a timely synthesis of the current research landscape on non-invasive technologies for pediatric stress detection and intervention. By examining 49 studies published over the past decade, we mapped out prevailing trends in sensing modalities, study populations, validation protocols, and intervention strategies.

To our knowledge, this is one of the first reviews to integrate both physiological and behavioral sensing frameworks within a pediatric population. In doing so, it reveals important advances—such as the emergence of hybrid sensing systems, caregiver-inclusive mobile platforms, and game-based interventions—while also surfacing clear limitations around dataset diversity, ecological deployment, ethical design, and cross-study comparability.

Despite progress, the field remains highly fragmented. Most systems continue to rely on short-term lab

deployments, with few examples of scalable or longitudinal monitoring in real-world contexts. Very few studies incorporate participatory co-design methods, even though child engagement and comfort are critical to success. Validation remains inconsistent, and open, multimodal, pediatric-specific datasets are still rare.

Based on our findings, we propose the following research and design priorities:

- Develop and standardize pediatric-specific stress induction and validation protocols.
- Broaden in-the-wild deployments across diverse cultural, home, and school contexts.
- Promote hybrid sensing approaches that balance signal fidelity and comfort.
- Encourage open sharing of ethically curated, multimodal datasets.
- Embed transparency, privacy safeguards, and co-design in system development.

This review underscores that stress detection in children is not just a technological challenge, but a human-centered, ethical, and contextual one. Progress will require sustained interdisciplinary collaboration across engineering, psychology, education, ethics, and design. Only then can we create robust, inclusive, and scalable solutions that genuinely support children's emotional well-being across the environments they grow, play, and learn in.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

This work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Generative AI Use

Generative AI tools (ChatGPT/GPT-5) were used only to assist with language editing and formatting. The authors reviewed and take full responsibility for the content of this manuscript. No AI tools were used for data extraction, analysis, or interpretation.

Authors' Contributions

AN conceived and designed the review. SE and MB conducted the searches, screening, and data charting. MB resolved screening disagreements and provided methodological oversight. AN drafted the manuscript; SE and MB critically revised the manuscript. All authors approved the final version.

Data Availability

No new human participant data were collected. All charted data and per-study extractions generated during this review are included in the article and its Multimedia Appendices (Appendix 1). Database-specific search strategies and exact queries are provided in Multimedia Appendix 2. The PRISMA-ScR checklist is provided in Multimedia Appendix 3.

Abbreviations

ASD : Autism Spectrum Disorder

EDA : Electrodermal Activity

EMA : Ecological Momentary Assessment

HR : Heart Rate

HRV : Heart Rate Variability

PPG : Photoplethysmography

rPPG : Remote Photoplethysmography

TSST : Trier Social Stress Test

mHealth : Mobile Health

Ethics Approval and Informed Consent

Not applicable. This scoping review synthesizes previously published studies and did not involve new human participants.

Multimedia Appendix 1. Per-Study Extraction Table

A CSV/LaTeX table with one row per included study (n=49), including: sample size and population type; age group; setting; stress induction; ground truth/validation; sensing modality and form factor; evaluation method(s) and metrics; mitigation features.

Multimedia Appendix 2. Database Search Strategies

For each database (PubMed, Scopus, IEEE Xplore, ACM Digital Library) provide platform, date of last search (June 2025), and the exact query strings (with field tags and limits). Include a note that Google Scholar was screened as a supplementary source.

Multimedia appendix 3. PRISMA-ScR Checklist

Completed PRISMA-ScR checklist indicating where each item is addressed in the manuscript (section/page references).

References

- [1] Yaribeygi H, Panahi Y, Sahraei H, Johnston TP, Sahebkar A. The impact of stress on body function: A review. EXCLI Journal 2017;16:1057–72. <https://doi.org/10.17179/excli2017-480>.
- [2] Kivimäki M, Steptoe A. Effects of stress on the development and progression of cardiovascular disease. Nature Reviews Cardiology 2018;15:215–29.
- [3] Smith KE, Pollak SD. Early life stress and development: Potential mechanisms for adverse

outcomes. *Journal of Neurodevelopmental Disorders* 2020;12:34. <https://doi.org/10.1186/s11689-020-09337-y>.

- [4] Hovsepian K, al'Absi M, Ertin E, Kamarck T, Nakajima M, Kumar S. cStress: Towards a gold standard for continuous stress assessment in the mobile environment. *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, 2015, p. 493–504.
- [5] Cohen S, Kamarck T, Mermelstein R. *Perceived stress scale*. Mind Garden; 1994.
- [6] National Institutes of Health. [Ethical issues in pediatric research](#) 2023.
- [7] Gedam S, Paul S. A review on mental stress detection using wearable sensors and machine learning techniques. *IEEE Access* 2021;9:84045–66.
- [8] Hickey BA et al. Smart devices and wearable technologies to detect and monitor mental health conditions and stress: A systematic review. *Sensors* 2021;21:3461.
- [9] Can YS et al. How to relax in stressful situations: A smart stress reduction system. *Healthcare* 2020;8:100.
- [10] Smith EN et al. Integrating wearables in stress management interventions: Promising evidence from a randomized trial. *International Journal of Stress Management* 2020;27:172.
- [11] Shterev K. [HRV-based stress detection](#) 2025.
- [12] González Ramírez ML et al. Wearables for stress management: A scoping review. *Healthcare* 2023;11:2369. <https://doi.org/10.3390/healthcare11172369>.
- [13] Arsalan A, Majid M, Nizami IF, Manzoor W, Anwar SM, Ryu J. Human stress assessment: A comprehensive review of methods using wearable sensors and non-wearable techniques. *arXiv Preprint arXiv:220203033* 2022.
- [14] Antle AN, McLaren ES, Fiedler H, Johnson N. Design for Mental Health: How Socio-Technological Processes Mediate Outcome Measures in a Field Study of a Wearable Anxiety App. *Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction*, New York, NY, USA: Association for Computing Machinery; 2019, p. 87–96. <https://doi.org/10.1145/3294109.3295650>.
- [15] MedTech News. *Wearable technology in pediatric healthcare: Challenges and future directions* 2023.
- [16] Mase A et al. Non-contact measurement of stress using microwave reflectometry. *IEEE sensors*, 2010, p. 2011–5.
- [17] Azevedo AI, Santos HD, Sá VJ, Lopes NV. Behavioural Biometrics for Authentication and Stress Detection – A Case Study with Children. In: Jahankhani H, Carlile A, Emm D, Hosseinian-Far A, Brown G, Sexton G, et al., editors. *Global Security, Safety and Sustainability - The Security Challenges of the Connected World*, Cham: Springer International Publishing; 2016, p. 266–80. https://doi.org/10.1007/978-3-319-51064-4_21.
- [18] Vijayan V et al. Deep learning methods for wearable sensor data in healthcare. *Sensors* 2023;23:512.
- [19] Zisopoulou O et al. Stress detection in dental anxiety using non-contact methods. *Healthcare Technology Letters* 2022;9:42–50.
- [20] Aqajari SA et al. [Context-aware stress monitoring](#) 2023.*ff*
- [22] Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology* 2018;18:143. <https://doi.org/10.1186/s12874-018-0611-x>.
- [23] Evans JStBT. [Deductive reasoning](#). In: Holyoak KJ, Morrison RG, editors. *The cambridge handbook of thinking and reasoning*, Cambridge University Press; 2005, p. 169–84.
- [24] Azungah T. Qualitative research: Deductive and inductive approaches to data analysis. *Qualitative Research Journal* 2018;18:383–400.
- [25] Yen P-Y, Bakken S. Review of health information technology usability study methodologies. *Journal of the American Medical Informatics Association* 2012;19:413–22.

<https://doi.org/10.1136/amiajnl-2010-000020>.

[26] Factor—Glossary F. Glossary of terms. 2023. Available from: https://csrc.nist.gov/glossary/term/form_factor. Accessed September 25, 2025.

[27] Deng L, Rattadilok P, Xiong R. A Machine Learning-Based Monitoring System for Attention and Stress Detection for Children with Autism Spectrum Disorders. Proceedings of the 2021 International Conference on Intelligent Medicine and Health, New York, NY, USA: Association for Computing Machinery; 2022, p. 23–9. <https://doi.org/10.1145/3484377.3484381>.

[28] Aktas SNB, Uluer P, Coskun B, Toprak E, Barkana DE, Kose H, et al. Stress Detection of Children With ASD Using Physiological Signals. 2022 30th Signal Processing and Communications Applications Conference (SIU), Safranbolu, Turkey: IEEE; 2022, p. 1–4. <https://doi.org/10.1109/SIU55565.2022.9864668>.

[29] Coşkun B, Ay S, Erol Barkana D, Bostancı H, Uzun İ, Oktay AB, et al. A physiological signal database of children with different special needs for stress recognition. Scientific Data 2023;10:382. <https://doi.org/10.1038/s41597-023-02272-2>.

[30] Kadirvelu B, Bellido Bel T, Wu X, Burmester V, Ananth S, Cabral CC Branco B, et al. Mindcraft, a mobile mental health monitoring platform for children and young people: Development and acceptability pilot study. JMIR Formative Research 2023;7:e44877.

[31] Buddi P, Prasad V, Sunitha KVN, Reddy N, Anil C. DetectStress: A Novel Stress Detection System Based on Smartphone and Wireless Physical Activity Tracker: AICC 2018, 2019, p. 67–80. https://doi.org/10.1007/978-981-13-1580-0_7.

[32] Chaowadee N, Lertsiriyothin P, Phuangkhemkhao T, Chanwimalueang T. Reinforced Learning in Children through a Stress Warning Unit. 2021 2nd Information Communication Technologies Conference (ICTC), 2021, p. 386–9. <https://doi.org/10.1109/ICTC51749.2021.9441507>.

[33] Choi Y, Jeon Y-M, Wang L, Kim K. A Biological Signal-Based Stress Monitoring Framework for Children Using Wearable Devices. Sensors 2017;17:1936. <https://doi.org/10.3390/s17091936>.

[34] Coşkun B, Barkana DE, Uzun İ, Bostancı H, Tarakçı D. Classification of Stress and Participation using Physiological Signals of Children During Serious Game-Based Therapy. 2023 Medical Technologies Congress (TIPTEKNO), 2023, p. 1–4. <https://doi.org/10.1109/TIPTEKNO59875.2023.10359219>.

[35] Coşkun B, Uluer P, Toprak E, Barkana DE, Kose H, Zorcec T, et al. Stress Detection of Children with Autism using Physiological Signals in Kaspar Robot-Based Intervention Studies. 2022 9th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob), 2022, p. 01–7. <https://doi.org/10.1109/BioRob52689.2022.9925485>.

[36] Jin C, Osotsi A, Oravec Z. Predicting Adolescent Female Stress with Wearable Device Data Using Machine and Deep Learning. 2023 IEEE 19th International Conference on Body Sensor Networks (BSN), 2023, p. 1–4. <https://doi.org/10.1109/BSN58485.2023.10331414>.

[37] Vural ŞF, Yurdusever B, Oktay AB, Uzun I. Stress recognition from facial images in children during physiotherapy with serious games. Expert Systems with Applications 2024;238:121837. <https://doi.org/10.1016/j.eswa.2023.121837>.

[38] Yu Z, Sherpa MT, Iadarola S, Shamlan K, Daley S, Levine G, et al. Understanding Stress in Children with ASD and Their Caregivers in Daily Life: A Feasibility Study Using Mobile Devices. Proceedings of the 25th International Conference on Mobile Human-Computer Interaction, New York, NY, USA: Association for Computing Machinery; 2023, p. 1–7. <https://doi.org/10.1145/3565066.3608686>.

[39] Isbister K, Cottrell P, Cecchet A, Dagan E, Theofanopoulou N, Bertran FA, et al. Design (Not) Lost in Translation: A Case Study of an Intimate-Space Socially Assistive “Robot” for Emotion Regulation. ACM Trans Comput-Hum Interact 2022;29:32:1–36. <https://doi.org/10.1145/3491083>.

[40] Lutin E, De Raedt W, Steyaert J, Van Hoof C, Evers K. Exploring the perception of stress in childhood and early adolescence. Journal of Experimental Child Psychology 2023;228:105604.

<https://doi.org/10.1016/j.jecp.2022.105604>.

- [41] Redd CB, Silvera-Tawil D, Hopp D, Zandberg D, Martiniuk A, Dietrich C, et al. Physiological Signal Monitoring for Identification of Emotional Dysregulation in Children. 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada: IEEE; 2020, p. 4273–7. <https://doi.org/10.1109/EMBC44109.2020.9176506>.
- [42] Zhang H, Cao L, Feng L, Yang M. Multi-modal interactive fusion method for detecting teenagers' psychological stress. *Journal of Biomedical Informatics* 2020;106:103427. <https://doi.org/10.1016/j.jbi.2020.103427>.
- [43] Goulart C, Valadão C, Delisle-Rodriguez D, Funayama D, Favarato A, Baldo G, et al. Visual and Thermal Image Processing for Facial Specific Landmark Detection to Infer Emotions in a Child-Robot Interaction. *Sensors* 2019;19:2844. <https://doi.org/10.3390/s19132844>.
- [44] Kumazaki H, Yoshikawa Y, Matsumoto Y, Miyao M, Ishiguro H, Muramatsu T, et al. An Intervention for Children with Social Anxiety and Autism Spectrum Disorders Using an Android Robot. In: Otake M, Kurahashi S, Ota Y, Satoh K, Bekki D, editors. *New Frontiers in Artificial Intelligence*, Cham: Springer International Publishing; 2017, p. 470–7. https://doi.org/10.1007/978-3-319-50953-2_34.
- [45] Loftness BC, Halvorson-Phelan J, O'Leary A, Cheney N, McGinnis EW, McGinnis RS. UVM KID Study: Identifying Multimodal Features and Optimizing Wearable Instrumentation to Detect Child Anxiety. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual International Conference* 2022;2022:1141–4. <https://doi.org/10.1109/EMBC48229.2022.9871090>.
- [46] Sarabadani S, Schudlo LC, Samadani AA, Kushski A. Physiological Detection of Affective States in Children with Autism Spectrum Disorder. *IEEE Transactions on Affective Computing* 2020;11:588–600. <https://doi.org/10.1109/TAFFC.2018.2820049>.
- [47] Carlier S, Van der Paelt S, Ongenae F, De Backere F, De Turck F. Using a Serious Game to Reduce Stress and Anxiety in Children with Autism Spectrum Disorder. *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, New York, NY, USA: Association for Computing Machinery; 2019, p. 452–61. <https://doi.org/10.1145/3329189.3329237>.
- [48] Kim T-Y, Měsíček L, Kim S-H. Modeling of Child Stress-State Identification Based on Biometric Information in Mobile Environment. *Mobile Information Systems* 2021;2021:1–3. <https://doi.org/10.1155/2021/5531770>.
- [49] Nicolaidou I, Tozzi F, Kindynis P, Panayiotou M, Antoniadis A. DEVELOPMENT AND USABILITY OF A GAMIFIED APP TO HELP CHILDREN MANAGE STRESS: AN EVALUATION STUDY. *Italian Journal of Educational Technology* 2019;27:105–20. <https://doi.org/10.17471/2499-4324/1050>.
- [50] Theofanopoulou N, Isbister K, Edbrooke-Childs J, Slovák P. A Smart Toy Intervention to Promote Emotion Regulation in Middle Childhood: Feasibility Study. *JMIR Mental Health* 2019;6:e14029. <https://doi.org/10.2196/14029>.
- [51] Masino AJ, Forsyth D, Nuske H, Herrington J, Pennington J, Kushleyeva Y, et al. M-Health and Autism: Recognizing Stress and Anxiety with Machine Learning and Wearables Data. 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), 2019, p. 714–9. <https://doi.org/10.1109/CBMS.2019.00144>.
- [52] Melinda M, Setiawan V, Yunidar Y, Gopal Sakarkar, Nurlida Basir. Portable Stress Detection System for Autistic Children Using Fuzzy Logic. *Jurnal Nasional Teknik Elektro* 2024;60–8. <https://doi.org/10.25077/jnte.v13n2.1203.2024>.
- [53] Nguyen J, Cardy RE, Anagnostou E, Brian J, Kushki A. Examining the effect of a wearable, anxiety detection technology on improving the awareness of anxiety signs in autism spectrum disorder: A pilot randomized controlled trial. *Molecular Autism* 2021;12:72.

<https://doi.org/10.1186/s13229-021-00477-z>.

- [54] Northrup CM, Lantz J, Hamlin T. Wearable Stress Sensors for Children With Autism Spectrum Disorder With In Situ Alerts to Caregivers via a Mobile Phone. *Iproceedings* 2016;2:e9. <https://doi.org/10.2196/iproc.6119>.
- [55] Payani M, Raja KM, Saraan J, Srihari P. Design and Implementation of Human Stress Detection System: Especially for Autistic children. *Journal of Physics: Conference Series* 2021;1979:012063. <https://doi.org/10.1088/1742-6596/1979/1/012063>.
- [56] Wang P, Huang L, Dai G, Li J, Hu J, Barakova E, et al. Evaluating the Role of Interactive Encouragement Prompts for Parents in Parent–Child Stress Management. *Applied Sciences* 2025;15:256. <https://doi.org/10.3390/app15010256>.
- [57] Yao C, Wang P, Li Z, Li J, Li X, Hu J, et al. StressButton: Using Connectivity of Everyday Objects for Children to Coping Learning Stress. 2023. <https://doi.org/10.54941/ahfe1003257>.
- [58] Chiang Y-C, Chien SH-L, Lyu J-L, Chang C-K. Recognition of Dynamic Emotional Expressions in Children and Adults and Its Associations with Empathy. *Sensors* 2024;24:4674. <https://doi.org/10.3390/s24144674>.
- [59] Li J, Barakova E, Hu J, Staal W, Dongen-Boomsma M van. ApEn: A stress-aware pen for children with autism spectrum disorder. In: Ferrández Vicente JM, Álvarez-Sánchez JR, Paz López F de la, Adeli H, editors. *Artificial intelligence in neuroscience: Affective analysis and health applications*, Cham: Springer International Publishing; 2022, p. 281–90.
- [60] Nag A, Haber N, Voss C, Tamura S, Daniels J, Ma J, et al. Toward Continuous Social Phenotyping: Analyzing Gaze Patterns in an Emotion Recognition Task for Children With Autism Through Wearable Smart Glasses. *Journal of Medical Internet Research* 2020;22:e13810. <https://doi.org/10.2196/13810>.
- [61] V P HV, Shahkutty A, S A. Deep Learning-Based Stress Identification in Children: Unraveling Physiological Patterns. *IJIREEICE* 2023;12. <https://doi.org/10.17148/IJIREEICE.2024.12103>.
- [62] Li J, Wang P, Barakova EI, Hu J, Dai G. Stress Diffuser: A Biofeedback Agent for Stress Management in Children During Homework with Parent Involvement. *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference*, Delft Netherlands: ACM; 2024, p. 701–5. <https://doi.org/10.1145/3628516.3659378>.
- [63] Curzio O, Billeci L, Belmonti V, Colantonio S, Cotrozzi L, De Pasquale CF, et al. Horticultural Therapy May Reduce Psychological and Physiological Stress in Adolescents with Anorexia Nervosa: A Pilot Study. *Nutrients* 2022;14:5198. <https://doi.org/10.3390/nu14245198>.
- [64] Gul Airij A, Bakhteri R, Khalil-Hani M. Smart wearable stress monitoring device for autistic children. *Jurnal Teknologi* 2016;78. <https://doi.org/10.11113/jt.v78.9453>.
- [65] Rathod M, Dalvi C, Kaur K, Patil S, Gite S, Kamat P, et al. Kids' Emotion Recognition Using Various Deep-Learning Models with Explainable AI. *Sensors* 2022;22:8066. <https://doi.org/10.3390/s22208066>.
- [66] Rodríguez-Pellejero JM, Mulero-Henríquez I, Santana Amador Z. Real-time stress monitoring in a child-friendly court: A repeated measures field study. *Humanities and Social Sciences Communications* 2024;11:913. <https://doi.org/10.1057/s41599-024-03410-w>.
- [67] Shonkoff JP, Boyce WT, Bush NR, Gunnar MR, Hensch TK, Levitt P, et al. Translating the Biology of Adversity and Resilience Into New Measures for Pediatric Practice. *Pediatrics* 2022;149:e2021054493. <https://doi.org/10.1542/peds.2021-054493>.
- [68] Singh S. [Meltdown detection in autistic children combining stress sensors and machine learning](#). Master's thesis. State University of New York at Oswego, 2022.
- [69] Thierfelder A, Primbs J, Severitt B, Hohnacker CS, Kühnhausen J, Alt AK, et al. Multimodal Sensor-Based Identification of Stress and Compulsive Actions in Children with Obsessive-Compulsive Disorder for Telemedical Treatment. 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2022, p. 2976–82.

<https://doi.org/10.1109/EMBC48229.2022.9871899>.

- [70] Vitullo K, Benitez M. A wearable therapy and technology garment for kids: The underlying super hero. Proceedings of the 23rd International Symposium on Wearable Computers, London United Kingdom: ACM; 2019, p. 329–33. <https://doi.org/10.1145/3341163.3346933>.
- [71] Xiao Y, Sharma H, Tumanova V, Salekin A. Psychophysiology-aided Perceptually Fluent Speech Analysis of Children Who Stutter. Proceedings of the ACM/IEEE 16th International Conference on Cyber-Physical Systems (with CPS-IoT Week 2025), 2025, p. 1–1. <https://doi.org/10.1145/3716550.3722019>.
- [72] Chyan P, Achmad A, Nurtanio I, Areni IS. STRESS DETECTION OF CHILDREN THROUGH SPEECH SIGNALS IN MULTI-SPEAKER ENVIRONMENT USING DEEP LEARNING. International Journal of Innovative Computing, Information and Control 2023;19:1983–94. <https://doi.org/10.24507/ijicic.19.06.1983>.
- [73] Gurusamy D, Janarthanan M, Chakrabarti P. Analysis of Schoolgirls' Mental Health – A Machine Learning Approach to Distinguish Between Academic and Abusive Stress. Universal Journal of Public Health 2024;12:136–49. <https://doi.org/10.13189/ujph.2024.120115>.
- [74] Wang P, Zhou Y, Dai G, Li Z, Zhao X, Lu J, et al. Exploring Perceptions of Children's Learning Stress for Stress Management. In: Arai K, editor. Intelligent Computing, Cham: Springer Nature Switzerland; 2024, p. 51–65. https://doi.org/10.1007/978-3-031-62281-6_4.
- [75] Schmidt P, Reiss A, Duerichen R, Marberger C, Laerhoven KV. Introducing WESAD, a multimodal dataset for wearable stress and affect detection. Proceedings of the 20th ACM international conference on multimodal interaction (ICMI), 2018, p. 400–8. <https://doi.org/10.1145/3242969.3242985>.
- [76] Shu L, Yang J, Zhang Y, Gao J, Liu P. StressID: A multimodal dataset for stress identification in children. Sensors 2022;22:5185. <https://doi.org/10.3390/s22145185>.
- [77] Yang M, Li Q, Wang L, al. et. MMSD: Multimodal mental state dataset of adolescents with physiological and behavioral signals. IEEE Access 2023;11:32458–72. <https://doi.org/10.1109/ACCESS.2023.3254082>.
- [78] Wang R, Chen F, Chen Z, Li T, al. et. StudentLife: Assessing mental health, academic performance and behavioral trends of college students using smartphones. Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing, 2014, p. 3–14. <https://doi.org/10.1145/2632048.2632054>.
- [79] Prajod S, Krishnan A, Arora A. HRV-based stress detection models: Cross-dataset generalization challenges and solutions. IEEE Journal of Biomedical and Health Informatics 2024. <https://doi.org/10.1109/JBHI.2024.3371256>.
- [80] Kotsakidis I, D'Mello S, Liu C, Ferrara E. [Harmonizing multimodal stress datasets with ComBat for improved cross-study generalization](#). IEEE Transactions on Affective Computing 2025.
- [81] Menéndez S, Bello Orgaz G. Data transparency and ethics in pediatric wearables. International Journal of Child Health and Human Development 2023;16:47–56.

Supplementary Files

Multimedia Appendixes

Per-study extraction table.

URL: <http://asset.jmir.pub/assets/8f9310309812e41b0b0c64d0b6c5b71e.xlsx>

Database search strategies.

URL: <http://asset.jmir.pub/assets/4b3e036b4534d5c8dd6ddb8113238567.docx>

RISMA-ScR checklist.

URL: <http://asset.jmir.pub/assets/58402f372c0013af635d8b467be4a145.docx>