

Pilot study of a real-time early agitation capture technology (REACT) for children with intellectual and developmental disabilities

Nibraas Khan¹ , Abigale Plunk², Zhaobo Zheng³, Deeksha Adiani¹, John Staubitz⁴, Amy Weitlauf⁴ and Nilanjan Sarkar^{1,2,3}

Abstract

Objective: Children and adolescents with intellectual and developmental disabilities (IDD), particularly those with autism spectrum disorder, are at increased risk of challenging behaviors such as self-injury, aggression, elopement, and property destruction. To mitigate these challenges, it is crucial to focus on early signs of distress that may lead to these behaviors. These early signs might not be visible to the human eye but could be detected by predictive machine learning (ML) models that utilizes real-time sensing. Current behavioral assessment practices lack such proactive predictive models. This study developed and pilot-tested real-time early agitation capture technology (REACT), a real-time multimodal ML model to detect early signs of distress, termed “agitations.” Integrating multimodal sensing, ML, and human expertise could make behavioral assessments for people with IDD safer and more efficient.

Methods: We leveraged wearable technology to collect behavioral and physiological data from three children with IDD aged 6 to 9 years. The effectiveness of the REACT system was measured using F1 score, assessing its usefulness at the time of agitation to 20s prior.

Results: The REACT system was able to detect agitations with an average F1 score of 78.69% at the time of agitation and 68.20% 20s prior.

Conclusion: The findings support the use of the REACT model for real-time, proactive detection of agitations in children with IDD. This approach not only improves the accuracy of detecting distress signals that are imperceptible to the human eye but also increases the window for timely intervention before behavioral escalation, thereby enhancing safety, well-being, and inclusion for this vulnerable population. We believe that such technological support system will enhance user autonomy, self-advocacy, and self-determination.

Keywords

Machine learning, wearable sensing, autism spectrum disorder, multimodal data, challenging behaviors

Submission date: 18 January 2024; Acceptance date: 10 September 2024

¹Department of Computer Science, Vanderbilt University, Nashville, Tennessee, USA

²Department of Electrical and Computer Engineering, Vanderbilt University, Nashville, Tennessee, USA

³Department of Mechanical Engineering, Vanderbilt University, Nashville, Tennessee, USA

⁴Treatment and Research Institute for Autism Spectrum Disorders, Vanderbilt University Medical Center, Nashville, Tennessee, USA

Corresponding author:

Nibraas Khan, Department of Computer Science, Vanderbilt University, Nashville, Tennessee, 37235, USA.
 Email: nibraas.a.khan@vanderbilt.edu

Introduction

Children and adolescents with intellectual and developmental disabilities (IDD) show increased risk of displaying challenging behaviors.¹ “Challenging” behavior (as used in the behavior analytic literature^{2,3}), sometimes referred to as “problem behavior” or “severe behavior,” is defined in this work as behavior that places individuals and caregivers at risk for injury, results in property damage, or displaces individuals from their homes and communities. Examples of challenging behaviors include self-injury, aggression toward others, elopement, and property destruction, as documented in literature.^{4,5} Up to 50% of people with IDD present with some form of challenging behaviors, with 5–10% developing more severe presentations.⁶ This is especially true for children and adolescents with autism spectrum disorder (ASD),^{7,8} two-thirds of whom may display challenging behaviors before adulthood.⁹

Challenging behaviors exact a serious toll on individuals and families as well as educational and healthcare systems. Youths who present with severe challenging behaviors frequently experience restrictive management techniques, such as seclusion and restraint, and have increased emergency room visits or psychiatric hospitalizations related to these behavior concerns.^{4,5,9–11} Persistent challenging behaviors can put children and their caregivers at risk of potential physical harm¹² and prevent children from learning new skills,¹³ excluding them from school services and community opportunities that could increase their quality of life.¹⁴ This situation presents a critical challenge for individuals exhibiting these behaviors, their families, and caregivers, making it a key area for the development of innovative technologies aimed at enhancing outcomes within the community.

Caregivers of children with IDD often observe a pattern of agitation or early signs that may precede challenging behaviors. Agitation can manifest in various forms, some of which can be visible to a human observer and others which are more difficult to detect. An agitated state may show up as increased restlessness, changes in facial expressions, repetitive movements, or shifts in vocal tone. Caregivers, who are experts on their children, become skilled at recognizing subtle behavioral cues that may not be immediately evident to others. However, they usually cannot detect physiological markers of behavioral escalation, like changes in heart rate (HR), sweating, and other autonomic reactions. These are less salient and require specialized equipment for detection.

Caregivers of children with IDD often observe a pattern of agitation or early signs that may precede challenging behaviors. Agitation can manifest in various forms, some of which can be visible to a human observer and others which are more difficult to detect. An agitated state may show up as increased restlessness, changes in facial expressions, repetitive movements, or shifts in vocal tone.

Physiological responses require instrumentation to detect, such as changes in HR, sweating, and other autonomic reactions. These are telling signs of agitation that can provide early warning before behaviors become more pronounced. Caregivers, through their sustained and attentive interactions with their children, become skilled at recognizing subtle behavioral cues that may not be immediately evident to others, although they find it challenging to discern the physiological signs, which are less salient and often require specialized equipment to monitor effectively.

Recognizing the earliest signs of agitation, including subtle physiologic changes, is crucial. This provides an opportunity for timely intervention, potentially averting escalation of agitation into more severe incidents. By understanding and responding to early agitation, caregivers can employ specific strategies or interventions tailored to the individual’s needs, thereby creating a safer and more supportive environment. This insight from parents and caregivers is invaluable, as it contributes to a better understanding of the child’s behavior patterns and aids professionals in developing more effective behavioral interventions.^{15–17}

Identifying agitation before it escalates represents a significant advancement in behavior management strategies by proactively attempting to prevent behavior escalation instead of responding after a behavior has already occurred. The ability to foresee these early indicators, leveraging cutting-edge machine learning (ML) algorithms and wearable sensor technology, enables an early detection framework. ML algorithms are particularly adept at discerning physiological indicators of agitation, such as HR and sweat response, which are not apparent to human observation.^{18–20} Utilizing sensors, we can measure these physiological responses, providing data that ML algorithms can analyze to offer a reliable and objective assessment. These algorithms excel not only in analyzing physiological signals but also in observable behaviors; they can detect patterns and nuances in behavior, ensuring a comprehensive understanding of both overt and subtle signs of agitation. Such early detection facilitates interventions at the nascent stages of agitation that could potentially avert the escalation of behaviors. Ultimately, this predictive model could lead to a safer, more nurturing environment, enhancing the quality of care and reducing the overall stress for both individuals and their support systems.²¹

The integration of innovative technological applications into the field of applied behavior analysis (ABA), particularly in the detection of agitation, presents a significant opportunity to enhance current practices. Board Certified Behavior Analysts (BCBAs[®]), who conduct behavioral assessments, may face challenges detecting subtle agitation behaviors and have little opportunity to predict these early warning signs that challenging behavior may follow. These behaviors can be elusive, sometimes going unnoticed by caregivers or misinterpreted due to their subtlety. Even

with expert human observers, several indicators of agitation, like nuanced physiological responses or minor changes in posture, can be difficult to detect. This challenge is not a reflection of the limitations of BCBA^s but rather an inherent constraint of human observation.²² While BCBA^s are skilled in manipulating environmental factors to initiate and reinforce target behaviors, many existing methods for doing so do not allow for the development of robust predictive models that account for the intricate aspects of client behavior and physiology. Developing such models could significantly enhance the effectiveness of interventions for clients with challenging behaviors.

Affective computing is a field that enables intelligent systems to recognize, infer, and interpret human emotions and mental and behavior states using various signals such as behavioral (e.g. motion and eye gaze) and physiological (e.g. HR and body temperature) signals.²³ Goodwin et al. demonstrated that changes in physiological signals can be used to predict imminent aggression for youths with ASD who are minimally verbal using HR and galvanic skin response (GSR), also known as electrodermal activity.²⁴ Imbiriba et al. improved the accuracy of this study using principal component analysis and a nonlinear kernel-based support vector machine classifier with the same physiological responses.²⁵ Sarabadani et al. used an ensemble of classifiers to detect affective states induced by high and low arousal stimuli of positive and negative valence or high and low arousal using GSR, electrocardiograms, respiration (RSP), and temperature.²⁶ Inferred states such as users' engagement using GSR, RSP, and photoplethysmography (PPG) signals were used in a virtual reality driving simulator to improve engagement with the systems.²⁷ Motion data captured from children with ASD using accelerometers have been used to recognize body movements (e.g. hand flapping, body rocking, etc.) reflecting agitation or frustration in classroom settings.^{28,29} Body-worn accelerometers have also been used to recognize self-injurious behavior in children with ASD.³⁰

In addition to the work presented on deciphering affective states based on unimodal signals such as physiology or body gestures exclusively, several studies have shown promise in using data from multimodal sources. Multimodal sources allow for inferences that cannot be obtained from a single sensor or whose quality exceeds that of an inference drawn from a single source. Amat et al. and Tauseef et al. developed a collaborative virtual environment to encourage and measure teamwork between autistic and neurotypical adults in a workplace setting using speech, eye gaze, and game controller data.^{31,32} Plunk et al. expanded on virtual teamwork-based collaborative activities with semisupervised behavior labeling.³³ A human-machine interaction system for dementia intervention has been developed with multimodal physiological data collected from its users.^{34,35} Adiani et al. developed a multimodal job interview simulator for training

autistic individuals.³⁶ Siddharth et al. developed a wearable multimodal biosensing system that can collect, synchronize, record, and transmit data from multiple biosensors: PPG, EEG, eye-gaze headset, body motion capture, GSR, etc.³⁷ Communication and coordination skills in children with ASD were assessed using multimodal signals including speech, gestures, and synchronized motion.³⁸ Eye gaze, video recordings, and EEG have been used for early detection of ASD in children.³⁹

While the prediction of agitation that precedes challenging behavior is underresearched, strides have been made in the adjacent field of forecasting challenging behaviors. However, existing technological approaches to predicting challenging behaviors have several limitations. First, existing ML models for predicting challenging behaviors were not developed by embedding BCBA^s during training data collection. A prediction model grounded in clinical insight using BCBA input as ground truth will likely have closer overlap with expert human observation, be more accurate and meaningful, and enhance acceptance within care systems. Additionally, using a structured measurement context facilitates faster sampling of valid data than unstructured sampling.⁴⁰ However, previous pioneering studies have required nine or more hours of unstructured observation to build a predictive model.²⁵ In one recent study, it was necessary for participants to aggress up to 130 times during observations for the model to be built.²⁵ Repeatedly eliciting aggressive behaviors not only demands considerable time but also poses risks to the well-being and safety of both participants and observers. This human cost for building the model may be too much for some. Although expected to be robust and reliable, multimodal models have been underexplored in this field, thereby offering opportunities for innovation and improvement.

To overcome these limitations, our prior research⁴¹ introduced a predictive multimodal framework designed to alert caregivers about agitation behaviors in children with ASD. It established a clinically grounded, multimodal model that anticipated agitations in children with ASD, leveraging technological advancements for more effective intervention. Predictive multimodal framework to alert caregivers of problem behaviors for children with ASD, PreMAC, used commercial and customized sensors including a Microsoft Kinect, an E4 wristband (Empatica Inc.), a custom wearable intelligent noninvasive gesture sensor (WINGS), and a specifically developed tablet application for recording direct observations of participant behavior. While PreMAC has proven effective in offline settings, it has not yet been applied in real-time scenarios, which is crucial for interventions to be most effective. To meaningfully help intervention and reduce risk, proactive real-time detection of agitations is necessary.

In this study, we extend the work of, PreMAC,⁴¹ by working closely with BCBA^s, psychologists, caregivers,

and engineers to refine, deploy, and evaluate an ML model to proactively predict agitation behaviors. Specifically, in a controlled laboratory environment, we carried out behavior assessments in collaboration with BCAs and caregivers, carefully evoking behaviors identified as signals of agitation by caregivers, ensuring the process was safe for all participants. Simultaneously, our multimodal system collected physiological (e.g. HR and GSR) and behavioral (e.g. body movement and vocalization) changes. After the assessment, we used this information to train a ML model to recognize the underlying physiological and behavioral changes that could indicate the imminent occurrence of agitations. Significantly, our approach not only identifies participant agitation as it occurs but anticipates episodes before their observable onset, marking a crucial step forward in preempting and effectively preventing escalation to dangerous behaviors.

Recent advancements in assistive technologies have increasingly focused on empowering individuals with disabilities by enhancing rather than diminishing their agency.^{42,43} These technologies are developed to support more autonomous and self-directed living, enabling individuals to perform daily activities more independently and engage fully in community life. Technologies that integrate multiple modalities, such as sensory inputs from wearable devices, environmental controls, and communication aids, have been particularly effective in enhancing the functionality and independence of people with disabilities.^{44,45} Our approach empowers users and caregivers to take proactive steps, thereby fostering a sense of control and improving overall well-being. This approach not only aligns with contemporary efforts to enhance the agency of people with disabilities through technology but also contributes to the broader goals of inclusion and empowerment within the disability community.

The primary contribution of this work is to present a Pilot Study of a real-time early agitation capture technology (REACT) to detect imminent occurrence of agitation in real time. This extends our earlier work, PreMAC, which only provided offline analysis capabilities. In addition, we further extended our previous work by introducing a new auditory modality to capture agitation associated with vocal tone and verbalization. We preliminarily validate our work through a two-visit training-testing study with three children with IDD with varying agitations.

Methods

Participants

All procedures were approved by the Institutional Review Board. All participants and their caregivers provided informed assent and consent. Less verbal participants were provided with visual schedules to facilitate assent discussions or allow for study discontinuation. Caregivers

were present for the duration of study sessions and asked to advocate for their children by letting study staff know if behaviors shown were unusually intense, distressed, or unsafe.

Participants were recruited from an existing clinical database that includes individuals with and without documented developmental disabilities (such as ASD) whose guardians have consented to be contacted about research opportunities. Eligibility criteria included being between 3–17 years of age; being from a family that spoke primarily English within the home; having a high likelihood of tolerating the wearable devices (no significant sensitivities to fabric textures, sleeve length, and so on); and presenting with predictable, frequent, and observable challenging behaviors that could be evoked by a therapist as part of a session. Interested caregivers completed an intake interview with a study BCBA to determine eligibility and appropriateness based on the child's phenotypic and behavior profile. The study's purpose and activities, including the use of wearable devices, were explained, with caregivers engaged in discussion about participant sensory profiles and the likelihood of tolerating such procedures before the initial study visit. Additionally, we sought to learn if the participants were likely to escalate relatively quickly with an unfamiliar adult in a clinical setting, and if so, would they have the same capacity to return to a calm state. If the caregiver and the BCBA agreed that participants could likely tolerate activities and engage safely in the research protocol, they were enrolled in the study.

At the study visit, participants and their caregivers completed written informed consent and assent through which study procedures were again described. Participants were given the opportunity to interact with and try on the wearable sensors, facilitating an understanding of the technology prior to participation. Ample time was provided for both participants and their representatives to ask questions and express any concerns.

In addition to the use of visual support for less verbal participants, observable dissenting behaviors were monitored and interpreted with the help of therapists and caregivers, ensuring that participants' nonverbal cues were respected throughout the study. Importantly, participants were informed and reminded that they could ask to stop participation at any time during the study, further reinforcing our commitment to ethical research practices and respect for individual autonomy. This approach underscores the concept of informed consent as an ongoing process throughout the study duration.

In this work, we present data collection, analysis, and modeling as three individual case studies. Individualized modeling is necessary as each participant had varying ranges of physiological and behavioral signals, as well as different agitations.

Participant 1 was a 6-year-old male with minimal spontaneous spoken language. Challenging behaviors of

concern to his family included screaming, pushing items away, hitting himself in the leg or the face, and falling to the floor. Signs of agitation identified by his family included unhappy sounds, an upset face, arm flapping and tensing, and walking away.

Participant 2 was a 9-year-old male with an extensive and expressive vocabulary. Challenging behaviors of concern to his family included elopement and aggression toward adults and children (hitting, kicking, biting). Signs of agitation identified by his family included a raised voice, running away, verbal threats, movements (balled up fists, sudden jumping, throwing his hands at his sides), and pulling items from another person's grasp.

Participant 3 was an 8-year-old female with strong receptive language but articulation errors that impacted her expressive communication skills. Challenging behaviors of concern to her family included hitting, kicking, screaming, and falling to the floor. Signals of agitation identified by her family included forcefully tensing her body, swinging her arms, clenching her fists, or wiggling her body in her seat; using a louder or angrier voice, or speaking more rapidly; changing her facial expression (knitting her eyebrows); pushing items away, or stating, "you are annoying."

Experimental design

In our study, we developed a framework that allowed for the therapist to safely evoke agitation behaviors within a laboratory environment, as identified by caregivers. This framework was inspired by the principles of functional analysis (FA) and incorporates aspects of the interview-informed synthesized contingency analysis (IISCA). FA, a foundational approach in ABA, methodically identifies the functions of challenging behaviors, while IISCA specifically emphasizes the critical role of caregiver insights.⁴⁶ Working in collaboration with BCBA and caregivers, we conducted behavior assessments with a strong emphasis on maintaining participant safety throughout the process. Within this framework, BCBA engage with caregivers through an open-ended interview to understand the specific environmental and behavioral contexts of individuals.⁴⁷ This understanding informs the selection and presentation of stimuli that can either evoke behavior, known as establishing operations (EOs),⁴⁸ or serve as reinforcing stimuli (SRs) to manage or modify the behavior. The BCBA methodically presents and escalates EO, closely monitoring for signs of agitation, before delivering SRs. This approach is tailored to each case, guided by the caregiver's real-time feedback communicated through an earbud. An independent data collector observes the assessments and uses a data collection application to record the conditions of the assessment and the behaviors of the participant. The process of EO and SR presentation repeats once the BCBA has verified that the participant's behavior has

returned to normal levels and can receive the reinforcers indicated by their caregivers. The data gathered were then instrumental in training our ML model to proactively predict agitation behaviors in real-time, leveraging the clinical expertise of BCBA to inform model development.

For the study, each participant takes part in two assessments following this framework, with each assessment lasting approximately an hour. In Assessment 1, data are collected for training the predictive model, and in Assessment 2, signs of agitations are predicted in real-time using the trained model. During each assessment, a data collection tool called the data collection controller, which synchronizes all modalities, is deployed for multimodal data collection. To validate those results, a second BCBA observes the visit to engage in traditional methods of direct behavior observation that served as the ground truth regarding the occurrence of agitation or challenging behavior. Once the multimodal data have been collected and synchronized, the data are then denoised and used to extract features. The features extracted from Assessment 1's data are then used to train a ML model using the ground truth labels. This model is then used to predict the agitation in Assessment 2. Figure 1(a) highlights the process of training the ML model, and Figure 1(b) highlights the process of proactively predicting agitation before these episodes were observable using the underlying physiological and behavioral signals. All procedures were reviewed by a team of expert BCBA as well as parents of children with ASD and approved by the Institutional Review Board.

Because physiological and behavior signals as well as signs of agitation vary between individuals, we will be presenting our results as three individual case studies. While each participant was scheduled for a one-and-a-half-hour session, the assessment length varied across participants due to circumstances such as different sensitivities to EO and SRs, varying lengths of time to return to a calm state, and time limitations of the caregivers. Specifically, the number of episodes featuring observable agitation evoked within the assessments varied as certain participants tended to be more sensitive to EO while others were not.

Materials

As seen in Figure 1, different system components work together to capture annotated participants' behavioral and physiological signals during the assessments. A wearable physiological sensor is used to capture several peripheral physiological signals, a custom-made movement detection sensor, WINGS, is used to capture the upper body pose, and a wireless microphone is used to capture the audio signal. Additionally, a tablet-based application, behavior data collection integrator (BDCI), is used by BCBA to record each occurrence of the participants' behavior in real-time. These components, in conjunction, lead to a

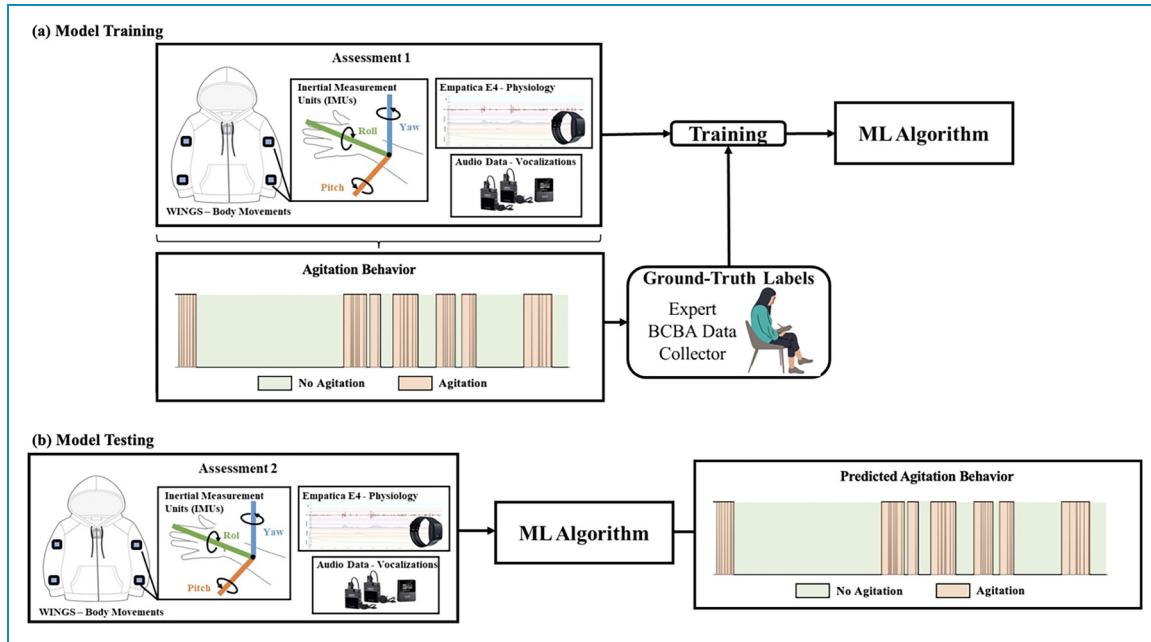


Figure 1. Two step training and testing mechanism: (a) schematic of training of the machine learning model using physiological and behavioral signals with BCBA provided identification of agitations; (b) schematic of the testing of the trained machine learning model using physiological and behavioral signals to predict the occurrence of agitations.
BCBA: Board Certified Behavior Analyst.

fully annotated dataset used for training and testing proactive real-time agitation detection ML models. Various technical components of our system are discussed below.

Wearable sensor for peripheral physiology. A medical-grade, noninvasive wearable physiological sensor called the E4 [<https://www.empatica.com/research/e4/>] is used to capture real-time blood volume pulse (BVP, 64 Hz) and GSR (64 Hz). The BVP signals are filtered through a 1 Hz to 8 Hz band-pass filter and used to calculate the HR. Additionally, tonic skin conductance level and phasic skin conductance response are extracted from the GSR signal⁴⁹ as they correlate with physiological arousal level.⁵⁰ Physiological signals are aggregated using a 60-s rolling window to account for their slow-moving nature.

WINGS for motion detection. We capture the upper body pose, such as hand flapping and body rocking, using the roll, pitch, and yaw of the upper limbs and torso of the participant through a custom-designed sensor called WINGS that uses five inertial measurement units (IMUs). The WINGS sensor, developed in our previous work,⁴¹ uses IMUs to capture bodily orientation in space using three internal sensors: an accelerometer, gyroscope, and magnetometer. By placing five IMUs, one on each shoulder, one on each elbow, and one on the back of the participants, the upper body pose can be reconstructed using the measurements from each sensor. Each IMU

wirelessly communicates its reading to an Arduino Uno microcontroller, which then transmits it to a computer workstation.⁵¹ The internal sensors of the IMU are used to compute the roll (θ), pitch (ψ), and yaw (ϕ) angles of the torso and limbs. With this information, the orientation of the participant's upper body can be calculated through the orientation of the shoulders, elbows, and back. The θ , ψ , and ϕ were computed using a one-second rolling window. Several methods exist to extract the θ , ψ , and ϕ , but we choose to use the following functions:⁵²

$$\theta = \tan^{-1} \left(\frac{accl_y}{\sqrt{accl_y^2 + accl_z^2}} \right) \quad (1)$$

$$\psi = \tan^{-1} \left(\frac{accl_x}{\sqrt{accl_y^2 + accl_z^2}} \right) \quad (2)$$

$$\phi = \frac{mag_z S\psi - mag_y C\theta}{\sqrt{mag_x C\theta + mag_y S\theta + mag_z C\psi S\theta}} \quad (3)$$

where $accl_x$, $accl_y$, and $accl_z$ are the accelerations in three directions, mag_x , mag_y , and mag_z are the strength of magnetic field in three directions, and S and C represent sine and cosine functions, respectively.

Microphone for audio signal. In this work, we also capture vocalizations, which have been shown to reflect at least

24 kinds of emotional states regardless of verbal content,⁵³ to detect agitations. These audio signals are captured from microphones and processed into various features such as pitch, volume, and zero-cross rate, which are then used for analysis. We use mel frequency cepstral coefficients (MFCCs), which are a compact representation of possibly infinite sinusoids in an audio signal.^{54–56}

We capture audio using a wireless lavalier lapel system. The system uses 2.4G wireless technology with dual transmitters and one receiver to wirelessly record audio for up to 5 h within a range of 164 feet sampled at 441 kHz. The raw audio signals are processed to generate the MFCCs. The MFCCs are extracted using six steps: (a) split the audio into small frames, (b) smoothen the transition between frames through windowing, (c) calculate the discrete Fourier transform, (d) apply 40 triangular filters known as filter banks, (e) pass through a log function, and (f) apply discrete cosine transform.⁵⁷

Behavior assessment application. We modified an application developed in a prior study that allows for behavior annotation using a reactive and responsive cross-platform application.⁵⁸ The application has three pages: initialization, assessment, and summary (Figure 2). The initialization page allows for the BCBA to input participant information, therapist information, assessment number, and assessment type (i.e. training or testing), as well as unique agitations and challenging behaviors for a specific participant. On the assessment page, the BCBA can record their observations

in real time while a JavaScript Object Notation file is produced with timestamped occurrences. Once a session has been completed, the summary page displays information pertinent to BCBA regarding the session.

Procedures

Figure 3 shows the experimental setup with two rooms, located in a clinical space, with a one-way mirror separating them. The participant and a BCBA are in the experimental room while the observers and additional BCBA are in the observational space. The participant wears an E4 sensor that collects physiological data. Participants are given the choice to wear it on either an ankle or a wrist. The WINGS system, which is a cotton hoodie outfitted with sensors, was worn on the upper body along with a lapel microphone placed in the hoodie. The BCBA in the experimental space is also wearing a lapel microphone. Both microphones send data to a receiver at a workstation in the observational space, which is also occupied by an engineering researcher, the participants' primary caregivers, the BCBA ground truth data collector, and a BCBA assessment manager. The BCBA assessment manager observes the assessment and takes input from the caregiver to guide the assessment toward a safe and data-rich outcome. The therapist in the experimental space communicates with the observers via Bluetooth earbud to ensure the time components of the experimental protocol are correctly executed.



Figure 2. Overview of the behavior assessment application, illustrating its three primary pages: initialization (for entering participant and session details), assessment (for real-time observation and data recording), and summary (displaying session insights postassessment).



Figure 3. Data collection setup, showing two connected spaces: the experimental room for participants and BCBA, equipped with sensors, and the observational room for the research team and caregivers, with equipment for monitoring and communication. BCBA: Board Certified Behavior Analyst.

Following a multielement design for single-subject research, the data collection experiment had both test (T) sessions, which alternated between EO and SR conditions, and control (C) sessions, which maintained SR conditions throughout. The typical session followed a Control-Test-Control-Test-Test (CTCTT) structure; however, this was occasionally modified during the assessment based on the BCBA's discretion. During EO presentations, the BCBA attempted to evoke agitation using conditions reported by the caregivers to lead to behavioral escalation. Commonly reported stimuli to evoke agitation included requesting the completion of homework assignments, removing preferred toys or electronics, and withdrawing preferred social attention. Following EO presentations in which agitation behaviors were witnessed, the BCBA provided preferred attention and items/activities for at least 90 s to allow for de-escalation during SR conditions. To ensure the participant had returned to a calm state (i.e. lack of agitation or challenging behavior), the caregiver provided feedback to both BCBA. Figure 4 provides the structure of the experimental procedure.

Caregivers always observed participants throughout each session and provided input on participant comfort and safety, as well as providing real-time feedback on the application of procedures intended to evoke escalation or calm states in the participant. Importantly, the goal was to develop a prediction model not for challenging behavior itself, but for agitations based on caregiver identification of signs of behavior escalation. That is, the goal was not to evoke unsafe behavior or prolong participant distress. Notably, no parent or participant requested to stop a session, underscoring the effectiveness of our approach in maintaining participant comfort and safety throughout the study.

Analysis

The predictive framework utilized Assessment 1 data for training a ML model, and Assessment 2 data for real-time testing and evaluation. During the assessments, the observing BCBA used the BDCI to note occurrences of

agitations. This created a log of agitation episodes with their corresponding timestamps that we used as ground-truth labels when training and evaluating the ML model. The following sections explain the steps the framework follows beginning with synchronizing and preprocessing the multimodal data.

Preprocessing. For our ML model, we used the following multimodal data inputs: physiological (E4) and behavior (WINGS and microphone). Due to the variability in sampling rate across modalities discussed in the Materials section, the collected data needed to be synchronized. In addition, features needed to be extracted to train the ML model. Features were extracted by applying a rolling window to the physiological, audio, and WINGS data after synchronization. The rolling window extracted the mean, minimum, and maximum values within the specified window. We chose these features based on our previous work⁴¹ as well as other published literature in the field^{21,37,59} that showed potential for these features to capture challenging behaviors. MFCCs were extracted from the audio modality. Table 1 shows the features extracted and their units. Table 2 details the length of each assessment and the number of agitations displayed in each assessment across the three participants.

ML model and real-time prediction. In our ML analysis, we utilized 32 ML models for proactive real-time prediction analysis such as Logistic Regression, AdaBoost Classifier, and Ridge Classifier. Each model was subjected to Bayesian optimization to determine the most effective hyperparameters through maximizing the F1 score, thereby ensuring optimal performance for each participant. Bayesian optimization is a powerful strategy for hyperparameter tuning, particularly effective in high-dimensional spaces. It works by constructing a probabilistic model that maps hyperparameters to a probability of a score on the objective function. This method iteratively updates the model with results from previous evaluations, using these insights to choose the next set of hyperparameters to evaluate. The models were implemented in Python using Scikit-Learn.⁶⁰ The optimization approach also included determining the ideal window length for our labels, acknowledging that agitations occur over a period rather than instantaneously; they start seconds before their annotation and last for a few seconds afterward. An example of a 30-s window is shown in Figure 5.

After a thorough evaluation, we chose the AdaBoost Classifier as our predictive model. Adaptive Boosting is an ensemble learning technique that works by combining multiple weak learners, typically decision trees, to create a robust classifier. Each successive model in the AdaBoost sequence focuses on the instances that were incorrectly predicted by the previous model, thereby

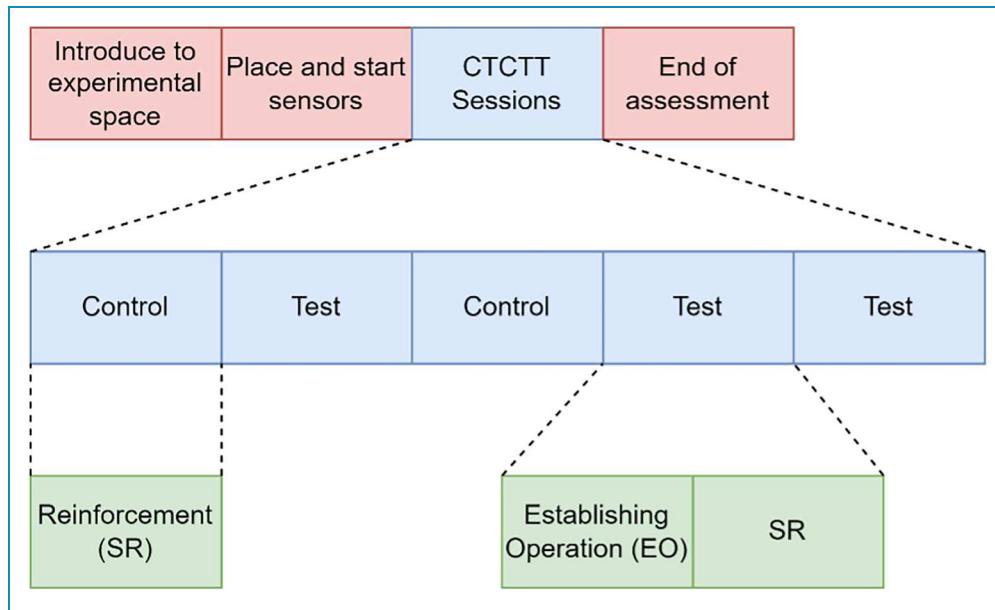


Figure 4. Overview of experimental procedure for BCBAs outlining the structured approach followed by BCBAs, a CTCTT sequence of test and control sessions, adaptable based on their discretion. It includes evoking agitations under specific conditions and implementing strategies for deescalation, with caregiver feedback ensuring participant calmness.

BCBA: Board Certified Behavior Analyst; CTCTT: Control-Test-Control-Test-Test.

Table 1. Features extracted for the machine learning model derived from physiological, audio, and WINGS data.

Modality	Features Extracted
Physiological data from E4	12 total features
	Average, standard deviation, minimum, maximum
	Heart rate (beats per minute)
	Skin conductance level (μ S)
	Skin conductance response (μ S)
Body movement data from WINGS	60 total features
	Average, standard deviation, minimum, maximum
	Roll, pitch, yaw (\circ)
Vocalization from microphone	12 total features
	MFCCs (unit of time)

MFCC: mel frequency cepstral coefficient; WING: wearable intelligent noninvasive gesture sensor.

continuously improving its performance. We trained the AdaBoost Classifier to predict agitation behaviors using the preprocessed data collected in Assessment 1. Once the

Table 2. Duration of each assessment and the frequency of agitations observed for each of the three participant case studies.

	Participant 1	Participant 2	Participant 3
Assessment 1 length	1 h 13 m	1 h 2 m	1 h 1 m
Assessment 1 number of agitations	18	18	16
Assessment 2 length	49 m	1 h 11 m	45 m
Assessment 1 number of agitations	14	21	13

model was trained with Assessment 1's data, Assessment 2's data were preprocessed to make real-time predictions by feeding the data into the model as if the data were coming from an experimental session as a real-time data stream. An additional step was taken to process the predictions made on Assessment 2's data before evaluating the performance of the model. Since an agitation is not an instantaneous event, we decided that if our model predicts an agitation within the identified window around an agitation, then the remainder of the band following the correct prediction should also be marked as correct as shown in Figure 6.

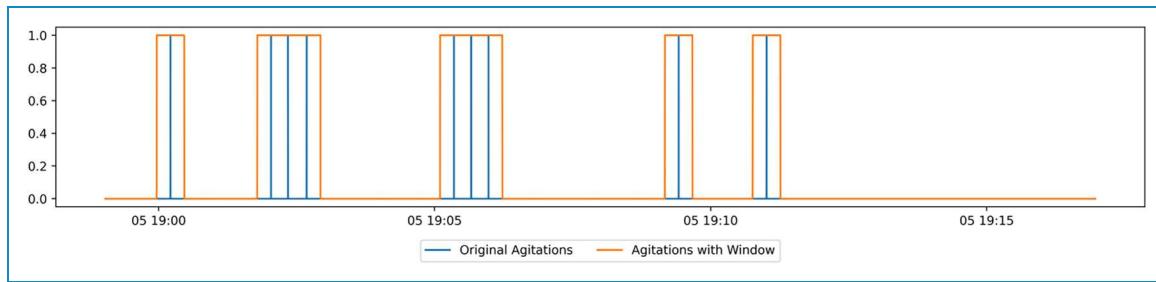


Figure 5. Example of applying a 30-s window to agitation labels, reflecting the duration over which agitations typically unfold.

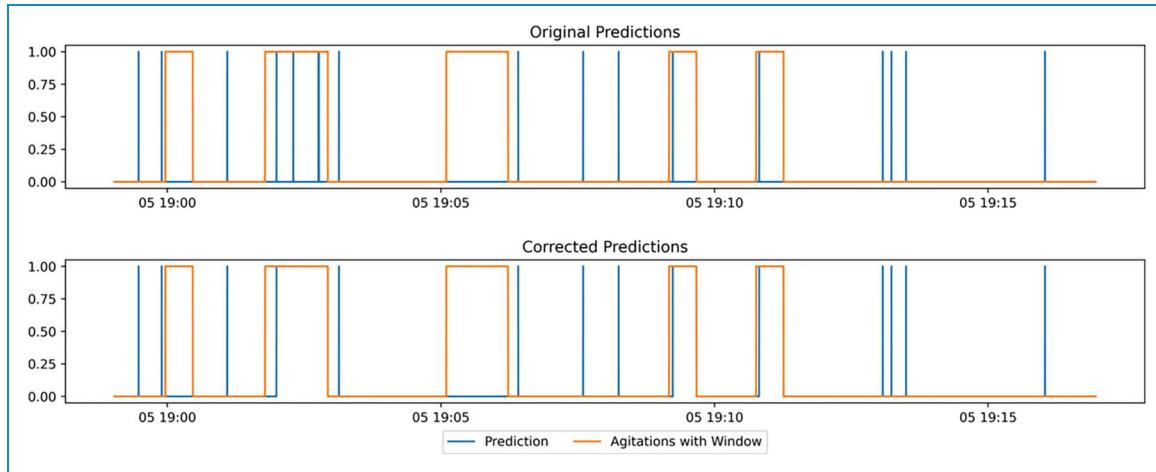


Figure 6. Example of how predictions were refined to account for the duration of agitation events, with a focus on marking the extended period following a correctly predicted agitation as accurate.

Model analysis. We evaluated the performance of our model using Assessment 2 data. Data from Assessment 2 were collected and fed into the model trained on Assessment 1 data as a real-time data stream to simulate prediction capability. To do so, we implemented a time shifts to simulate real-time predictions. We implemented six time shifts of 5 s, 10 s, 15 s, 20 s, 25 s, and 30 s to see how far in advance agitations can be predicted. Considering the specific requirements of our study, we emphasized the importance of the F1 score as a key metric in our model evaluation as shown in equation (4). It is vital in situations like ours where both false negatives and false positives carry significant consequences as our predictions need to be trusted. Missing an agitation (false negative) could potentially lead to challenging behaviors, and inaccurately predicting an agitation (false positive) could lead to mistrust in the model. This approach ensures that the model is not only sensitive to detecting agitations but also maintains a low rate of false predictions, striking a balance for effective and reliable predictions.

$$f1 = 2 \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

where precision is defined as the ratio of true positives (TP) to the sum of true positives and false positives (TP + FP), and recall is defined as the ratio of true positives to the sum of true positives and false negatives (TP + FN).

While recall, precision, and accuracy are important during analysis, they are not the primary metrics in this study. High recall ensures minimal missed detections of agitation precursors. However, recall alone could increase the propensity for false positives, which could undermine confidence in the model. Precision is valuable in confirming the relevance of detected behaviors, but relying on it exclusively does not capture the model's ability to consistently identify all relevant agitation precursors. Accuracy, reflecting the overall rate of correct predictions, does not always present a clear picture, especially in datasets where agitation events are sparse compared to nonagitation events.

Results

In our analysis, we evaluated numerous ML models. Table 3 presents the top four algorithms we identified based on performance, detailing the F1 score, recall, and

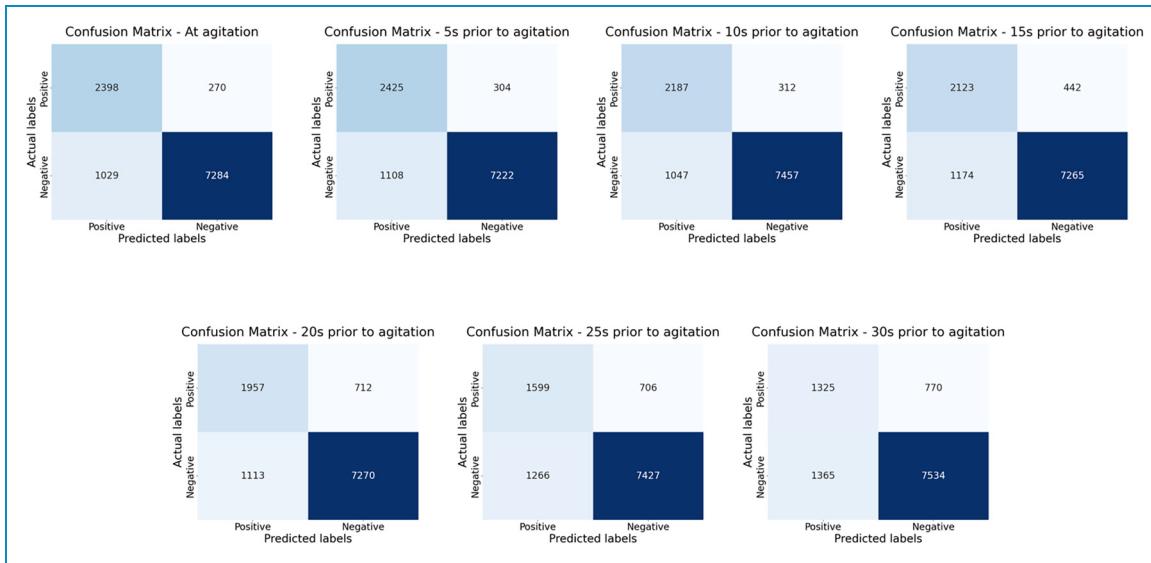
Table 3. Performance evaluation of the top four machine learning algorithms, presenting F1 score, recall, and accuracy metrics, complemented by their respective optimized hyperparameters.

	AdaBoost	Logistic regression	Random forest	K-neighbors
Hyper parameters	Estimator: Decision tree <i>n</i> estimators: 100 Learning rate: 1.0	Max iter: 2500	<i>n</i> estimators: 350 Max depth: 20	<i>n</i> neighbors: 15 Algorithm: kd tree Leaf size: 45
<i>At agitation</i>				
F1 score	78.69	78.15	77.96	77.91
Recall	69.97	68.70	69.50	69.41
Accuracy	88.17	88.04	87.99	87.53
<i>5 s prior to agitation</i>				
F1 score	77.45	77.10	76.36	76.69
Recall	68.64	68.27	67.84	68.23
Accuracy	87.23	87.44	87.54	86.78
<i>10 s prior to agitation</i>				
F1 score	76.30	75.94	76.25	75.12
Recall	67.63	67.77	66.47	66.62
Accuracy	87.64	86.49	87.30	87.29
<i>15 s prior to agitation</i>				
F1 score	72.43	71.28	71.63	72.09
Recall	64.39	64.38	63.88	63.23
Accuracy	85.31	84.63	84.82	85.21
<i>20 s prior to agitation</i>				
F1 score	68.20	67.90	68.13	67.25
Recall	63.75	63.75	62.78	63.16
Accuracy	83.89	82.70	83.92	82.77
<i>25 s prior to agitation</i>				
F1 score	61.86	60.86	61.19	61.58
Recall	55.81	55.13	54.79	54.97
Accuracy	82.07	81.59	82.19	81.35

(continued)

Table 3. Continued.

	AdaBoost	Logistic regression	Random forest	K-neighbors
30 s prior to agitation				
F1 score	55.38	54.25	54.64	55.09
Recall	49.26	48.72	49.03	48.90
Accuracy	80.58	80.39	79.43	79.49

**Figure 7.** Confusion matrices displaying the true positives, true negatives, false positives, and false negatives for the AdaBoost model used in agitation prediction. Each matrix corresponds to various distances from the agitation.

accuracy metrics, as well as the optimized hyperparameters for each. Among these, AdaBoost Classifier consistently demonstrated exceptional performance across the metrics, leading us to adopt it as our primary classifier for further analyses. Figure 7 displays the confusion matrices for the AdaBoost Classifier.

Figure 8 shows the metrics for each participant with different prediction distances. In general, we would expect the scores to drop as predictions are made further in advance of the agitation until a certain threshold. Across all three experiments, the scores began to drop below an acceptable performance when making predictions more than 20 s in advance. Our results showed an average F1 score across the three participants of 78.69% at the time of agitation, 77.45% at 5 s prior, 76.30% at 10 s prior, 72.43% at 15 s prior, 68.20% at 20 s prior, 61.86% at 25 s prior, and 55.38% at 30 s prior.

Additionally, we conducted an ablation study focusing on predictions at the time of agitations to gauge the significance of individual modalities on the overall performance of our classification model. In this study, we sequentially

omitted one of the following modalities: WINGS, HR, GSR, and audio. The objective was to discern the distinct contributions of each modality to the classifier's accuracy and probe for potential dependencies or redundancies. The outcomes of this ablation assessment provided valuable insights into the hierarchical significance of these modalities within our classification framework, as delineated in Table 4.

In our commitment to advancing scientific knowledge and supporting reproducibility, both the code and the dataset used in this study are available upon request through Vanderbilt University, via the following link: <https://lab.vanderbilt.edu/rasl/dataset-and-or-code-request/>.

Discussion

In our work, we employed ML and affective computing techniques to proactively identify agitations in children with IDD, as reported by caregivers. This approach is based on the understanding that these agitations, if

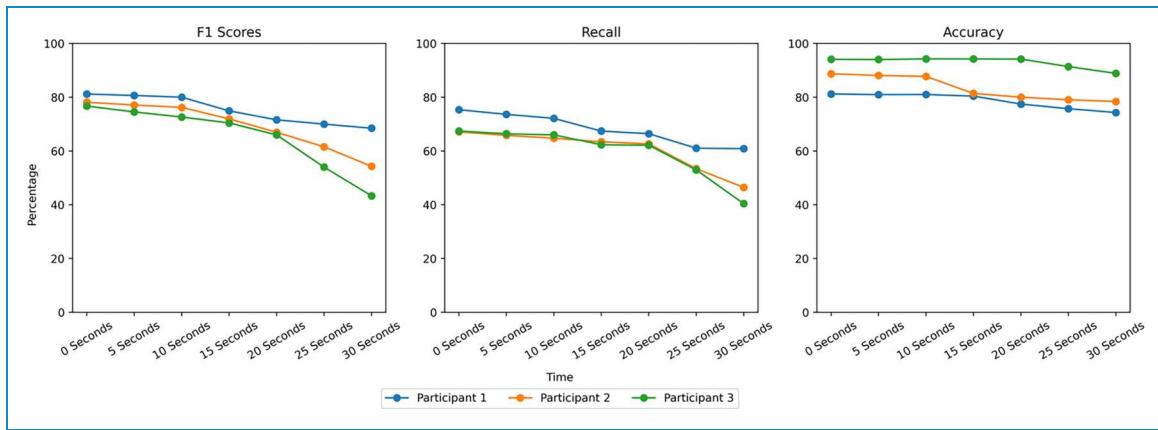


Figure 8. Comparative metrics across participants focusing on the change in performance metrics as predictions are made increasingly earlier before the agitation occurs.

Table 4. Ablation study results illustrating the impact on classifier performance when omitting individual modalities: WINGS, HR, GSR, and audio. This provides insights into each modality's contribution and potential interdependencies within the classification framework.

Excluded modalities	F1 score	Recall	Accuracy
None	78.69	69.97	88.17
WINGS	68.48	62.91	80.43
HR	71.87	64.03	81.94
GSR	69.13	60.94	81.21
Audio	71.34	63.32	83.41

GSR: galvanic skin response; HR: heart rate; WING: wearable intelligent noninvasive gesture sensor.

recognized promptly, could help in preventing their escalation into challenging behaviors.

We build upon existing work to demonstrate several promising technological advancements. First, we advanced the capabilities of our previous wearable system⁴¹ by integrating the capacity for proactive real-time prediction, enabling BCBA to adapt their interventions before participants' behavior escalates. Additionally, we expanded our multimodal data capture to include audio, to integrate vocalizations. This led to more robust agitation detection than achieved in our previous study. We also expanded upon our data capture application to include customizable, BCBA- defined agitations for child-specific annotation. Each of these achievements represents promising steps toward wearable systems capable of multimodal data capture and agitation detection within community settings.

Our preliminary trial with three participants highlighted the predictive value of our ML model. We were able to

predict agitation based on physiological, body movement behavior, and vocalization behavior data captured through various sensors. Our sensors accurately captured subtle signals such as HR and electrodermal activity, which are imperceptible to the observing BCBA. Sensors also captured joint angles of the upper body and vocalizations, both of which were described as important reflections of agitations by participants' caregivers. Body motions such as hand flapping and body rocking were recognized in other studies to detect agitation and self- injurious behaviors in children with ASD.^{28–30} Nonlexical vocal sounds and vocalization have been captured to recognize at least 24 kinds of various emotions.⁵³ Our work incorporates multimodal data through joint angles for body motion detection and vocalization for emotion to generate robust agitation detection models.

We found that as participants approached an occurrence of agitation, the signals we captured varied from their calm state. In this work, our algorithm capitalized on that deviation to determine that agitation was likely. Each participant had an individualized model trained during the first visit. These models were tested with data from the second visit. It was shown that agitation can be reliably predicted with an average F1 score of 68.20% at 20 s prior to the agitation behavior. While the predictive performance decreases as the prediction window extends further into the future, this is an expected tradeoff. The ability to provide earlier warnings, even with somewhat reduced performance, is a significant advantage, offering caregivers the opportunity to take preventative measures. Some forms of agitation are more readily detectable than others. For instance, behaviors like fist banging present clearer signals for the model, whereas more subtle actions such as a making a fist pose a greater challenge. Although these variations influence the overall predictive scores, the results are still promising. Despite the variation in detectability, the

models maintain a significant capacity to predict crucial signs of agitation, providing valuable insights for caregivers.

In our ablation study, we examined the distinct contributions of each modality by selectively omitting them and observing the resultant classification performance. The removal of WINGS led to discernible reductions in all metrics, underscoring its central significance. When we removed HR and GSR modalities, distinct outcomes emerged. Excluding HR resulted in a significant but less pronounced drop than the other modalities, highlighting its auxiliary role in the classification. In contrast, the omission of GSR led to performance metrics closely mirroring those seen with the removal of WINGS, elucidating its integral contribution. Notably, when audio data was excluded, the drop in metrics was more substantial than that observed with HR, emphasizing audio's pivotal role in the classification process. These results accentuate the individual and collective relevance of the modalities in enhancing classification outcomes.

To our knowledge, our study is the first to employ multimodal data capture and ML for the specific purpose of early-stage agitation prediction, shifting the focus from behaviors that have already escalated. Other teams have predicted imminent aggression (escalated behavior) one minute before they occur.²⁴ We take this concept further by focusing on early signs of agitation, as identified by caregivers, thereby reducing risk, and providing valuable time for preventative measures before behaviors escalate to challenging levels. Specifically, our goal is to maximize utility and participant safety during behavior assessment approaches that can place individuals at significant risk of harm. An important focus of ongoing work is evaluation not only of efficacy, but also of safety, tolerability, and acceptability to future potential users, including youths with challenging behaviors.

Building on the capability to proactively detect agitations, the next progression in behavioral analysis is the identification of precursors to problem behaviors. While agitation indicates a general state of distress—often identified by caregivers—it is the precursors that bear a closer relation to the challenging behaviors themselves. These precursors are the final observable behaviors before a challenge occurs, serving as a critical alert to imminent escalation. By focusing on these precursors, caregivers and clinicians are afforded crucial additional time to implement targeted interventions aimed at preventing the challenging behaviors from occurring.²¹ Within this context, our model has the potential to be an asset in FA and practical functional assessment (PFA). FA is a methodical approach for identifying the causes and functions of challenging behaviors through observation and manipulation of environmental variables for the purpose of treating challenging behavior. PFA, a novel set of FA procedures streamlines this process by focusing

on the immediate precursors to challenging behavior, thus minimizing the occurrence of challenging behaviors during the assessment. As we continue to refine our model, we anticipate it could play a critical role in the future of behavior assessments, offering a more proactive and less intrusive means to support individuals with IDD and their caregivers.

Ethical considerations

Introducing ML into clinical settings can open the door to privacy and security concerns. The REACT system involves the collection of sensitive behavioral data. Ensuring the privacy and security of this data is of utmost importance. Therefore, data collection and storage policies were clearly established, and any applications were encrypted to ensure the protection of privacy for participants.

An additional issue is ensuring informed consent and assent. Many individuals with IDD receiving services are nonspeaking, making it difficult to explain the wearable sensors and ensure comfort during data collection. Therefore, we set clear guidelines to ensure no one was forced to wear sensors against their will, to confirm understanding of the system by guardians, to respect the autonomy of the individuals, and to guarantee complete transparency when including ML in behavioral therapy.

ML is highly susceptible to bias. When developing algorithms, this is an important consideration especially when utilizing group models. One way to avoid this is to always utilize individualized ML models for agitation prediction as we did in this study. However, this is not always possible as training a model requires large amounts of data. When group models are utilized, it is essential to ensure equal representation across race, gender, and ethnicity to reduce bias in ML. In the application of predicting agitations, models must be trained on data representative of varying types of agitations as they are individual-specific.

The most crucial ethical consideration when introducing ML into a clinical setting is to acknowledge that while ML can offer valuable insights, it is not infallible and may occasionally generate errors or inaccurate recommendations. ML should be utilized to inform and aid clinicians in decision-making, not replace them. Therefore, it is essential for clinicians to exercise caution and never solely rely on ML without critical evaluation and human judgment.

It is crucial to emphasize that the REACT system is designed to augment, not replace, the human elements of care and intervention strategies. The application of ML in this context is intended to support BCBAs and caregivers by providing additional tools for understanding and responding to the needs of individuals with IDD, rather than diminishing the caregiver's role or the individual's autonomy. To ensure that technology is used respectfully and ethically, the REACT system must be integrated into care strategies in a way that promotes collaboration

between the technology and human caregivers. This approach fosters a balance on both technological and human insights, ensuring that interventions remain person-centered. Additionally, continuous training for caregivers on the ethical use of this technology and ongoing dialogue with stakeholders about its impact are vital in maintaining a humane approach to behavioral management. These steps help safeguard against any potential depersonalization and ensure that technology serves as an aid, enhancing the ability to deliver compassionate and individualized care.

Furthermore, incorporating these technologies into behavioral interventions not only augments the capabilities of BCAs but also promotes user autonomy and supports self-advocacy and self-determination among individuals with disabilities. By providing real-time data and predictive insights, these tools could empower individuals to better understand and communicate their own emotional and physiological states, fostering a greater sense of control over their environment and interactions. This shift toward technology-enhanced, personalized interventions aligns with contemporary goals in disability rights, emphasizing respect for individual agency, and the enabling of people with disabilities to lead more self-directed lives.

Limitations and future directions

Several limitations exist that warrant attention in future work. Individualized models were developed in our current study to account for the unique characteristics and variations in physiological and behavioral signals, as well as agitations, among participants. This personalized approach allowed us to tailor the predictive model to each individual's specific needs. However, as we continue our research, we envision moving toward the development of group models. By working with larger participant cohorts, we aim to explore the potential for building models that can generalize across individuals while still accommodating individual differences. The development of group models will enhance the scalability and broader application of our approach, facilitating more widespread implementation and benefiting a larger population of individuals with IDD.

We focused on creating individualized model in this initial investigative work to assess the feasibility of the presented approach. Although the results are promising, more participants will be required to generalize the findings of our approach across a broader spectrum of the IDD population. However, the insights gained from this limited cohort are critical for initial model testing, providing a foundational understanding for subsequent larger-scale studies. Future research will aim to include a larger number of participants, which will not only validate and potentially enhance the accuracy of our predictive models but also ensure that our findings are more broadly applicable and

representative of the diverse needs and conditions within the IDD community.

Our main goal during the assessments is to safely collect information about the most prevalent agitations to a child without pushing them toward challenging behavior. While this ensures safety, we are limited in how many occurrences of agitations can be evoked during assessments leading to a paucity of data for each agitation. Therefore, in future iterations of this work, longer assessments, or multiple assessments, particularly those already being conducted as behavior standards of care in the community, are required to collect more thorough training data.

Additionally, to reduce excessive stimulus from wires and electronics, the next version of WINGS will be wireless, and the clothing which holds WINGS needs to be purposefully built using comfortable clothing for children with various sensory profiles as determined through stakeholder input. We also found that the wristband used to capture physiological signals was prone to occasional loss of connectivity, dropping of data packets, and stagnant signals despite the high cost, warranting the need for the development of more reliable wearable physiological sensors. To make REACT portable, data collection and prediction need to be moved from computer-based to smartphone-based.

In this study, formal feedback from participants or their caregivers was not systematically collected as part of the methodology. However, it is important to note that during the study period, no adverse comments or concerns were raised by either caregivers or participants regarding the use of the technology. Recognizing the value of such feedback in evaluating the impact and acceptance of assistive technologies, future works will include a more structured feedback collection process in future studies. This will enable us to better understand the reception of the technology by the disability community and to ensure that it not only meets functional expectations but also aligns with the personal and social contexts of the users.

Despite these limitations, our current system can produce a robust ML algorithm for the proactive real-time detection of agitations with multimodal data collected using wearable sensors and a customized tablet for data annotation. In this work, we produced a real-time agitation prediction system to provide BCAs with valuable intervention time without escalating to challenging behaviors. Technology of this kind could improve the safety and efficiency of treatment protocols by notifying interventionists, caregivers, or, in the future, people with challenging behaviors themselves that behavior escalation is likely, prompting them to use evidence-based interventions for timely deescalation or redirection.

Conclusion

This study demonstrates the efficacy of REACT in proactively predicting early signs of agitation in children

with IDD using a multimodal data approach. By integrating ML techniques with sensory data capture from wearable technologies, our model successfully provides real-time, anticipatory alerts that allow caregivers and behavior analysts to intervene before behaviors escalate. The inclusion of multiple data modalities, such as physiological signals, body movements, and vocalizations, enriches the robustness of our predictions, significantly enhancing the predictive value of our model. Furthermore, our ablation study highlights the essential contributions of each modality to the overall performance, emphasizing the need for a comprehensive approach in future designs. As we continue to refine our system, we aim to extend its application to broader community settings and integrate it into existing behavioral analysis frameworks to maximize both participant safety and the effectiveness of interventions. The findings support the use of the REACT model for real-time, proactive detection of agitations in children with IDD. This approach not only improves the accuracy of detecting distress signals that are imperceptible to the human eye but also increases the window for timely intervention before behavioral escalation, thereby enhancing safety, well-being, and inclusion for this vulnerable population.

As development continues, maintaining an ongoing dialogue with the IDD community is crucial to ensure that the technology aligns with the real-world needs and preferences of those it is designed to support.

This collaboration is essential not only for optimizing functionality and effectiveness but also for reinforcing ethical standards that protect and empower the community of individuals with IDD. Through these efforts, technological advancements in behavioral analysis can continue to promote inclusion, respect, and empowerment for all individuals.

Acknowledgments: We thank Dr. Adithyan Rajaraman (Vanderbilt University Medical Center, Treatment and Research Institute for Autism Spectrum Disorders) for providing comments on the manuscript. We thank Lauren Shibley, BCBA, David Reichley, BCBA, and Amy Swanson, M.A. (Vanderbilt University Medical Center, Treatment and Research Institute for Autism Spectrum Disorders) for their contributions to these pilot experimental sessions. Additionally, we would like to thank all the children and their families for participating in the study.

Contributions: NK: methodology, analysis, investigation, data curation. AP: methodology, analysis, investigation, data curation. ZZ: conceptualization, investigation, funding acquisition. DA: analysis. JS: conceptualization, investigation. AW: conceptualization, supervision, project administration, funding acquisition. NS: conceptualization, resources, supervision, project administration, funding acquisition.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval: This study was approved by the Vanderbilt Institutional Review Board (211846), and written informed consent was obtained from legally authorized representative of all participants.

Funding: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Science Foundation (grant number 2124002).

Guarantor: NK

ORCID iD: Nibraas Khan  <https://orcid.org/0000-0002-2962-3119>

References

1. McClintock K, Hall S and Oliver C. Risk markers associated with challenging behaviours in people with intellectual disabilities: a meta-analytic study. *J Intellect Disabil Res* 2003; 47: 405–416.
2. Kurtz PF, Leoni M and Hagopian LP. Behavioral approaches to assessment and early intervention for severe problem behavior in intellectual and developmental disabilities. *Pediatr Clin N Am* 2020; 67: 499–511.
3. Newcomb ET and Hagopian LP. Treatment of severe problem behaviour in children with autism spectrum disorder and intellectual disabilities. *Int Rev Psychiatry* 2018; 30: 96–109.
4. Hyman SL, Fisher W, Mercugliano M, et al. Children with self-injurious behavior. *Pediatrics* 1990; 85: 437–441.
5. Thompson RH and Iwata BA. A comparison of outcomes from descriptive and functional analyses of problem behavior. *J Appl Behav Anal* 2007; 40: 333–338.
6. Dekker MC, Koot HM, van der Ende J, et al. Emotional and behavioral problems in children and adolescents with and without intellectual disability. *J Child Psychol Psychiatry* 2002; 43: 1087–1098.
7. Farmer CA and Aman MG. Aggressive behavior in a sample of children with autism spectrum disorders. *Res Autism Spectr Disord* 2011; 5: 317–323.
8. Soke GN, Rosenberg SA, Hamman RF, et al. Brief report: prevalence of self-injurious behaviors among children with autism spectrum disorder—a population-based study. *J Autism Dev Disord* 2016; 46: 3607–3614.
9. Kanne SM and Mazurek MO. Aggression in children and adolescents with asd: prevalence and risk factors. *J Autism Dev Disord* 2011; 41: 926–937.
10. Matson JL and Rivet TT. Characteristics of challenging behaviours in adults with autistic disorder, pdd-nos, and intellectual disability. *J Intellect Dev Dis* 2008; 33: 323–329.
11. Wisner-Carlson R, Uram S and Flis T. The transition to adulthood for young people with autism spectrum disorder. *Child Adolesc Psychiatr Clin N Am* 2020; 29: 345–358.
12. Chadwick O, Walker N, Bernard S, et al. Factors affecting the risk of behaviour problems in children with severe intellectual disability. *J Intellect Disabil Res* 2001; 44: 108–123.

13. Emerson E. *Challenging behaviour: analysis and intervention in people with severe intellectual disabilities*. Cambridge, England: Cambridge University Press, 2001.
14. Parish SL, Thomas KC, Rose R, et al. State medicaid spending and financial burden of families raising children with autism. *Intellect Dev Disabil* 2012; 50: 441–451.
15. Edelson SM. Understanding challenging behaviors in autism spectrum disorder: a multi-component, interdisciplinary model. *J Pers Med* 2022; 12: 1127.
16. Lane BR, Paynter J and Sharman R. Parent and teacher ratings of adaptive and challenging behaviours in young children with autism spectrum disorders. *Res Autism Spectr Disord* 2013; 7: 1196–1203.
17. Tathgur MK and Kang HK. Challenges of the caregivers in managing a child with autism spectrum disorder—a qualitative analysis. *Indian J Psychol Med* 2021; 43: 416–421.
18. Chaturvedi V, Kaur AB, Varshney V, et al. Music mood and human emotion recognition based on physiological signals: systematic review. *Multimedia Syst* 2022; 28: 21–44.
19. Ragot M, Martin N, Em S, et al. Emotion recognition using physiological signals: laboratory vs. wearable sensors. In: *AHFE 2017 international conference on advances in human factors and wearable technologies* (eds T Ahram and C Falcão), The Westin Bonaventure Hotel, Los Angeles, California, USA, 17 July–21 July 2017, pp.15–22. Cham, Switzerland: Springer International Publishing.
20. Saganowski S, Perz B, Polak AG, et al. Emotion recognition for everyday life using physiological signals from wearables: a systematic literature review. *IEEE Trans Affect Comput* 2023; 14: 1876–1897.
21. Warner CA, Hanley GP, Landa RK, et al. Toward accurate inferences of response class membership. *J Appl Behav Anal* 2019; 53: jaba.598.
22. Jessel J, Hanley GP and Ghaemmaghami M. On the standardization of the functional analysis. *Behav Anal Pract* 2020; 13: 205–216.
23. Picard RW. *Affective computing*. Cambridge, MA, USA: MIT Press, 2000.
24. Goodwin MS, Özdenizci O, Cumpanasou C, et al. Predicting imminent aggression onset in minimally-verbal youth with autism spectrum disorder using preceding physiological signals. In: *PervasiveHealth '18: Proceedings of the 12th EAI international conference on pervasive computing technologies for healthcare*, 2018, pp.201–207.
25. Imbiriba T, Cumpanasou DC, Heathers J, et al. Biosensor prediction of aggression in youth with autism using kernel-based methods. In: *Proceedings of the 13th ACM international conference on PErvasive technologies related to assistive environments (PETRA '20)*, Corfu, Greece, 2020, paper no: Article No. 13, pp.1–6. New York, NY, USA: Association for Computing Machinery (ACM).
26. Sarabadani S, Schudlo LC, Samadani AA, et al. Physiological detection of affective states in children with autism spectrum disorder. *IEEE Trans Affect Comput* 2020; 11: 588–600.
27. Bian D, Wade J, Swanson A, et al. Design of a physiology-based adaptive virtual reality driving platform for individuals with asd. *ACM Trans Access Comput* 2019; 12: 1–24.
28. Albinali F, Goodwin MS and Intille S. Detecting stereotypical motor movements in the classroom using accelerometry and pattern recognition algorithms. *Pervasive Mob Comput* 2012; 8: 103–114.
29. Rad NM, Bizzego A, Kia SM, et al. Convolutional neural network for stereotypical motor movement detection in autism. <https://arxiv.org/abs/1511.01865> (2015, accessed Sept 2024).
30. Cantin-Garside KD, Kong Z, White SW, et al. Detecting and classifying self-injurious behavior in autism spectrum disorder using machine learning techniques. *J Autism Dev Disord* 2020; 50: 4039–4052.
31. Amat AZ, Breen M, Hunt S, et al. Collaborative virtual environment to encourage teamwork in autistic adults in workplace settings. In: *Universal access in human-computer interaction. Design methods and user experience (HCII 2021)* (eds M Antona and C Stephanidis), Virtual, 2021, pp.339–348. Cham, Switzerland: Springer.
32. Tauseef M, Amat AZ, Adiani D, et al. Design of a virtual task to understand the nature of collaboration between autistic and neurotypical adults in workplace using multimodal data. In: *Universal access in human-computer interaction. User and context diversity (HCII 2022)* (eds M Antona and C Stephanidis), Virtual, 2022, pp.410–426. Cham, Switzerland: Springer.
33. Plunk A, Amat AZ, Tauseef M, et al. Semi-supervised behavior labeling using multimodal data during virtual teamwork-based collaborative activities. *Sensors* 2023; 23. Article 7.
34. Ghosh R, Khan N, Migovich M, et al. Iterative user centered design of robot-mediated paired activities for older adults with mild cognitive impairment (MCI). In: *Human aspects of IT for the aged population. Technology in everyday living (HCII 2022)* (eds Q Gao and J Zhou), Virtual, 2022, pp.14–28. Cham, Switzerland: Springer.
35. Migovich M, Ghosh R, Khan N, et al. System architecture and user interface design for a human-machine interaction system for dementia intervention. In: *Human aspects of IT for the aged population. Supporting everyday life activities (HCII 2021)* (eds Q Gao and J Zhou), Virtual, 2021, vol. 13331, pp.277–292. Cham, Switzerland: Springer.
36. Adiani D, Breen M, Migovich M, et al. Multimodal job interview simulator for training of autistic individuals. *Assist Technol* 2023; 36: 22–39.
37. Siddharth PA, Jung TP, et al. A wearable multi-modal biosensing system towards real-world applications. *IEEE Trans Biomed Eng* 2019; 66: 1137–1147.
38. Delaherche E, Chetouani M, Bigouret F, et al. Assessment of the communicative and coordination skills of children with autism spectrum disorders and typically developing children using social signal processing. *Res Autism Spectr Disord* 2013; 7: 741–756.
39. Liao M, Duan H and Wang G. Application of machine learning techniques to detect the children with autism spectrum disorder. *J Healthc Eng* 2022; 2022: 1–10.
40. Tiger JH, Miller SJ, Mevers JL, et al. On the representativeness of behavior observation samples in classrooms. *J Appl Behav Anal* 2013; 46: 424–435.
41. Zheng ZK, Staibitz JE, Weitlauf AS, et al. A predictive multimodal framework to alert caregivers of problem behaviors for children with asd (premac). *Sensors* 2021; 21: 370.
42. Dange P, Mistry T and Mann S. AI and assistive technologies for persons with disabilities-worldwide trends in the scientific

production using bibliometrix r tool. In: *International conference on artificial intelligence: towards sustainable intelligence* (eds S Tiwari, F Ortiz-Rodríguez, S Mishra, et al.), Pune, India, 4 September-5 September 2023, pp.24–43. Cham, Switzerland: Springer.

43. Muharib R, Ledbetter-Cho K, Bross LA, et al. Handheld technology to support vocational skills of individuals with intellectual and developmental disabilities in authentic settings: a systematic review. *J Autism Dev Disord* 2022; 9: 1–12.
44. Wehmeyer ML, Davies DK, Stock SE, et al. Applied cognitive technologies to support the autonomy of people with intellectual and developmental disabilities. *Adv Neurodev Disord* 2020; 4: 389–399.
45. Bhatti I, Tariq M, Hayat Y, et al. A multimodal affect recognition adaptive learning system for individuals with intellectual disabilities. *Eur J Sci Innov Technol* 2023; 3: 346–355.
46. Hanley GP, Jin CS, Vanselow NR, et al. Producing meaningful improvements in problem behavior of children with autism via synthesized analyses and treatments. *J Appl Behav Anal* 2014; 47: 16–36.
47. Hanley GP. Functional assessment of problem behavior: dispelling myths, overcoming implementation obstacles, and developing new lore. *Behav Anal Pract* 2012; 5: 54–72.
48. Laraway S, Snyderski S, Michael J, et al. Motivating operations and terms to describe them: some further refinements. *J Appl Behav Anal* 2003; 36: 407–414.
49. Lim CL, Rennie C, Barry RJ, et al. Decomposing skin conductance into tonic and phasic components. *Int J Psychophysiol* 1997; 25: 97–109.
50. Setz C, Arnrich B, Schumm J, et al. Discriminating stress from cognitive load using a wearable eda device. *IEEE Trans Inf Technol Biomed* 2010; 14: 410–417.
51. Warren ZE, Zheng Z, Swanson AR, et al. Can robotic interaction improve joint attention skills? *J Autism Dev Disord* 2015; 45: 3726–3734.
52. Evans DJ. On the representation of orientation space. *Mol Phys* 1977; 34: 317–325.
53. Cowen AS, Elfenbein HA, Laukka P, et al. Mapping 24 emotions conveyed by brief human vocalization. *Am Psychol* 2019; 74: 698–712.
54. Hidayat R and Winursito A. A modified mfcc for improved wavelet-based denoising on robust speech recognition. *Int J Intell Eng Syst* 2021; 14: 12–21.
55. Mohanta A and Mittal VK. Analysis and classification of speech sounds of children with autism spectrum disorder using acoustic features. *Comput Speech Lang* 2022; 72: 101287.
56. Poornima S and Kousalya G. Deep learning based behavioral analysis and exploration of emotions in ASD children. In: *2022 second international conference on artificial intelligence and smart energy (ICAIS)*, Coimbatore, India, 23 February-25 February 2022, pp.504–509. Piscataway, New Jersey: IEEE (Institute of Electrical and Electronics Engineers).
57. Sahidullah M and Saha G. Design, analysis and experimental evaluation of block based transformation in MFCC computation for speaker recognition. *Speech Commun* 2012; 54: 543–565.
58. Khan N, Ghosh R, Migovich M, et al. A data collection and annotation tool for asynchronous multimodal data during human-computer interactions. In Gao Q and Zhou J (eds) *Human aspects of IT for the aged population. Design, interaction and technology acceptance*. Cham: Springer International Publishing, 2022, pp.201–211.
59. Zheng ZK, Bian D, Swanson A, et al. Soft-brush: a novel tendon driven tactile stimulator for affective touch in children with autism. In: *Advances in human factors and ergonomics in healthcare and medical devices (AHFE 2018)* (ed N Lightner), Orlando, Florida, USA, 2018, pp.15–22. Cham, Switzerland: Springer.
60. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in python. *J Mach Learn Res* 2011; 12: 2825–2830.