UVM KID Study: Identifying Multimodal Features and Optimizing Wearable Instrumentation to Detect Child Anxiety

Bryn C. Loftness[1], *Student Member, IEEE,* Julia Halvorson-Phelan[2], Aisling O'Leary[1], Nick Cheney[1], Ellen W. McGinnis[2], Ryan S. McGinnis[1], *Senior Member, IEEE*

Abstract— **Anxiety and depression, collectively known as internalizing disorders, begin as early as the preschool years and impact nearly 1 out of every 5 children. Left undiagnosed and untreated, childhood internalizing disorders predict later health problems including substance abuse, development of comorbid psychopathology, increased risk for suicide, and substantial functional impairment. Current diagnostic procedures require access to clinical experts, take considerable time to complete, and inherently assume that child symptoms are observable by caregivers. Multi-modal wearable sensors may enable development of rapid point-of-care diagnostics that address these challenges. Building on our prior work, here we present an assessment battery for the development of a digital phenotype for internalizing disorders in young children and an early feasibility case study of multi-modal wearable sensor data from two participants, one of whom has been clinically diagnosed with an internalizing disorder. Results lend support that sacral movement responses and R-R interval during a short stress-induction task may facilitate child diagnosis. Multi-modal sensors measuring movement and surface biopotentials of the chest and trapezius are also shown to have significant redundancy, introducing the potential for sensor optimization moving forward. Future work aims to further optimize sensor placement, signals, features, and assessments to enable deployment in clinical practice.**

Clinical Relevance— **This work considers the development and optimization of technologies for improving the identification of children with internalizing disorders.**

Keywords— **digital medicine, anxiety, depression, wearable sensors, signal processing, pediatric mental health**

I. Introduction

Childhood internalizing disorders (anxiety and depression) are common, impairing, and have the potential to disrupt development well into adulthood [1]–[3]. Children 8 years and younger are at heightened risk for being overlooked, because they cannot reliably report their own emotional suffering [4]. Current screening tools include lengthy parent-report surveys which inherently under-report symptoms as child thoughts and emotions are difficult to identify even by adults who know the children best [5], [6]. Given these barriers and limited access to expert clinical assessment in most communities [7], there is an urgent need for objective, accurate, and low-burden tools for detecting anxiety and depression in young children.

In our previous work [8]–[13], we presented preliminary data indicating that digital phenotyping could enable detection of early childhood internalizing disorders. Movement and speech data were collected via a single belt-worn wearable sensor and a video camera, respectively, as the child engaged in one of four brief behavioral tasks intended to induce anxiety, fear, or pleasure according to NIMH Research Domain Criteria (RDoC) criteria [14]. Our current on-going study builds on that work by considering additional physiological measures, including skin temperature, electromyography (EMG), electrodermal activity (EDA), and electrocardiography (ECG) across multiple behavioral tasks to improve screening accuracy, and to examine multiple forms and locations of wearables to address barriers in deploying this approach for broad research and clinical use.

Herein, we introduce the KID (Kiddie Internalizing Disorder) Study and examine preliminary data of two biomarkers, movement and R-R interval, during a fear task for candidate features that may detect children with internalizing psychopathology. We also identify redundant sensor signals that point to future reductions in sensor array complexity that will simplify deployment. The overarching goal of this work is to demonstrate feasibility and potential directions for future work in pre-school aged children.

II. Methods

Participants (children aged 4-8 years old, and their caregiver) were recruited from community advertisements across the state of Vermont, over-selected for children with elevated anxiety screening scores. Involvement included one 3-hour laboratory visit. After consenting, the child was outfitted with wearable sensors and brought to a separate room while the caregiver was administered a gold-standard semi-structured diagnostic interview about their child's mental health (KSADS-PL), and completed surveys regarding demographics and mental health on both their child and themselves. Simultaneously, the child engaged in the behavioral assessment battery comprised of four mood induction tasks which represent RDoC Positive (Bubbles and Reward tasks) and Negative (Approach and Speech tasks) Valence domains that underly a breadth of internalizing disorders [15], [16]. Mood induction tasks are consistent with activities that children conduct during their everyday lives

*Research supported by the United States National Science Foundation (NSF Award #2046440) and National Institutes of Health (NIH Award MH123031)

Authors are affiliated with the University of Vermont (UVM)[1] or UVM Medical Center[2], Burlington, VT 05405 USA (corresponding author e-mail: Bryn.Loftness@uvm.edu).

playing with siblings or in school-settings. Rest periods, for physiological recovery, occur before and after each task.

1. Approach Task (30 seconds): Induces anxiety in young children [8] via potential threat. Child is led into a novel, dimly lit room toward an unknown and hidden object.
2. Speech Task (3 minutes): Adapted version of the Trier Social Stress Task for children (TSST-C [9]). Child asked to tell a 3-minute story that will be judged based on how interesting it is. An unexpected buzzer interrupts the child at 90 and 150 seconds. Following the task, children are given positive feedback.
3. Bubbles Task (3 minutes): Adapted version of the LAB-TAB bubble task meant to elicit positive affect [17] via hedonic response to reward. Child plays with bubbles from a bubble machine for 3 minutes. Administrator gives positive reflective feedback (i.e., "This is fun!").
4. Reward Task (1 minute): Adapted version of the delay task [18]. Child is offered a small, desirable toy after a delay of 1 minute to induce anticipatory positive affect.

After completing the behavioral assessment battery, children are administered an IQ test (Differential Abilities Scale), and free play until their caregiver has completed all surveys. Families are compensated for their participation. Caregivers are also given the results of the diagnostic interview after consensus diagnosis has been discussed.

Prior to assessment, the child is outfitted with five MC10 BioStamps, an Empatica E4, and a smartphone equipped with a custom data collection app (Fig. 1). BioStamps placed on the trapezius (trap), chest, and extensor digitorum record surface biopotential (EMG/ECG), accelerometer, and gyroscope data. BioStamps placed on the thigh and lower back (sacrum) record just accelerometer and gyroscope data. Study activities were approved by the local institutional review board.

III. Results

Herein, we present BioStamp data from two of the four tasks (Approach and Speech) and consider child diagnostic status from the gold-standard diagnostic interview. In so doing, we identify candidate signal features, not available in our prior work [9], which may aid in the detection of internalizing disorders in young children. We also identify information common across BioStamp sensors which could point to future reductions in sensors required for deployment.

A. Candidate Diagnostic Features

Figure 2 presents sacral movement and R-R interval data collected during the Speech Task from two children in the KID Study: one child affected by multiple internalizing disorders (Female, 8.5 years old, Diagnosed Generalized Anxiety Disorder, Separation Anxiety Disorder, Specific Phobia Disorder) and one child unaffected by any internalizing disorders (Male, 5.5 years old).

Sacral movement, as characterized by angular velocity magnitude (AVM), was different between the affected and unaffected children during the Speech Task. While the time series data (Fig. 2, top panel) indicates that the amplitude and frequency of movements are similar throughout the task, the 5 seconds following each of the unexpected buzzers (vertical cyan lines, Fig. 2) reveal significant differences. The affected child appears to respond more significantly to the buzzers (Fig. 2, middle panel) as indicated by larger sacral angular velocity magnitudes (Mann-Whitney U-test $p<0.001$ for both periods). While sacrum data are reported here, these same relationships were observed in data from the trap and chest. Interestingly, experimenter annotations after watching a recording of the Speech Tasks indicate that neither child had an observable movement response to the buzzers. Thus, the wearable sensors may be picking up on easily unnoticed, subtle movements.

We also derived R-R intervals from the BioStamp chest ECG signal (via [19]). Median R-R interval from a baseline period at the start of the behavioral assessment battery was used to normalize the Speech task R-R interval values for each child (Fig. 2, bottom). If a child's heart rate was higher than baseline during the Speech Task, we would expect values less than zero (shorter R-R interval). Both the unaffected child and affected child had R-R intervals that were significantly below zero (Wilcoxon signed-rank, $p<.001$), and thus elevated heart rates were observed during the Speech task. However, the median difference from baseline for the affected child is significantly larger than for the unaffected child (0.062 vs 0.016 seconds).

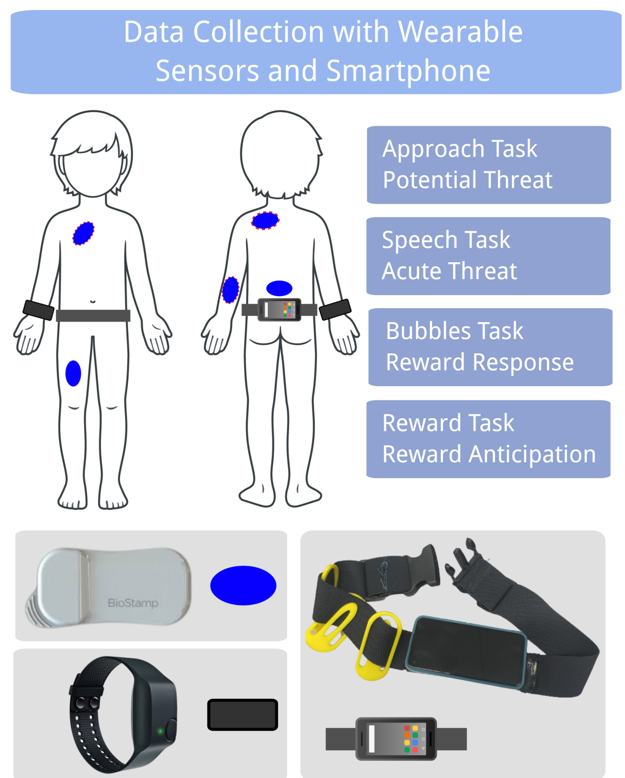

Figure 1. Three types of multi-modal sensors are attached to each child: 5 MC10 BioStamps; 1 Empatica E4 wrist device; and 1 smartphone attached via a custom elastic waistbelt. The BioStamps that measure surface biopotential, accelerometer, and gyroscope data are indicated in blue with a red dotted outline. The BioStamps that measure just accelerometer and gyroscope data are indicated in blue without an outline.

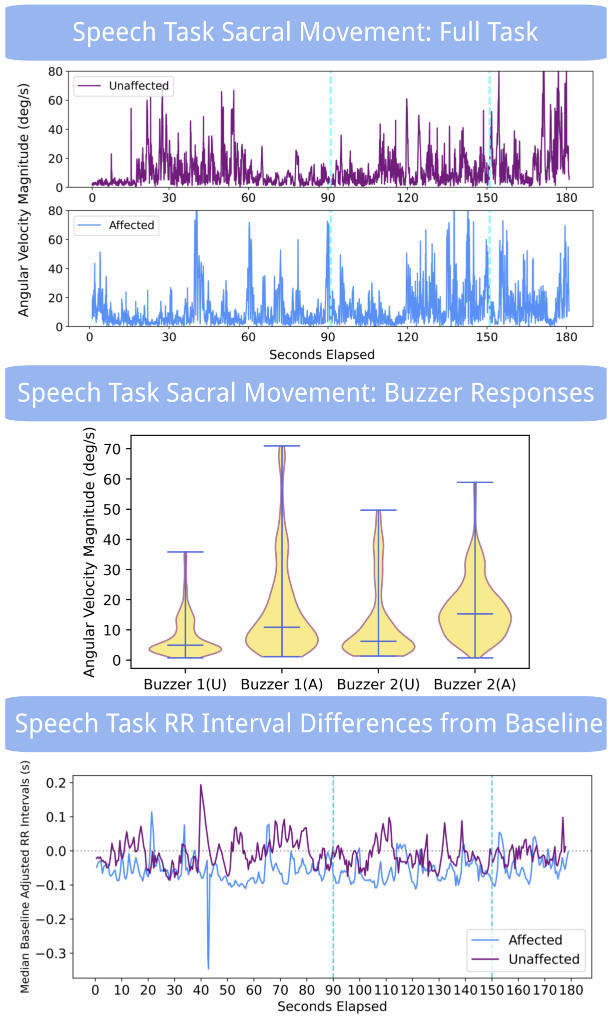

Figure 2. Timeseries angular velocity magnitude (top panel) for the unaffected child (purple) and affected child (blue) with timing of the surprising buzzers indicated with cyan vertical lines. Significant differences in amplitude of movement response to each buzzer were identified between children (middle panel). Baseline adjusted ECG-derived RR-interval timeseries (bottom panel) from the chest BioStamp also exhibit differences between children during the Speech Task.

B. Sensor Redundancy

Figure 3 presents examples of movement and biopotential data recorded from the trap and chest sensors during the Speech (R-R) and Approach (movement) Tasks. The angular velocity magnitude from the chest and trap sensors is highly correlated (Fig. 3, middle and bottom left, Pearson r=0.96, p<0.001). Interestingly, the surface biopotential signals from these device locations both exhibited clear R-waves that could be used for computing instantaneous heart rate and heart rate variability. To this end, the R-R interval time series extracted from each device are also highly correlated (Fig. 3, top and bottom right, Pearson r=0.74, p<0.001). Remaining differences in the R-R timeseries between sensors are likely due to corruption of the ECG signal from the trap sensor by muscle contractions. These can be reduced considerably with appropriate signal processing in future work [20].

IV. Discussion

In this study, we demonstrate the feasibility of using wearable instrumentation during a behavioral assessment battery for young children (4-8 years) to identify candidate digital features for detecting internalizing disorders and examine instrumentation optimization. We have extended previous work by collecting data from multiple sensor modalities and body locations. These data have never been collected synchronously in children, and particularly with patch-based wearable sensors.

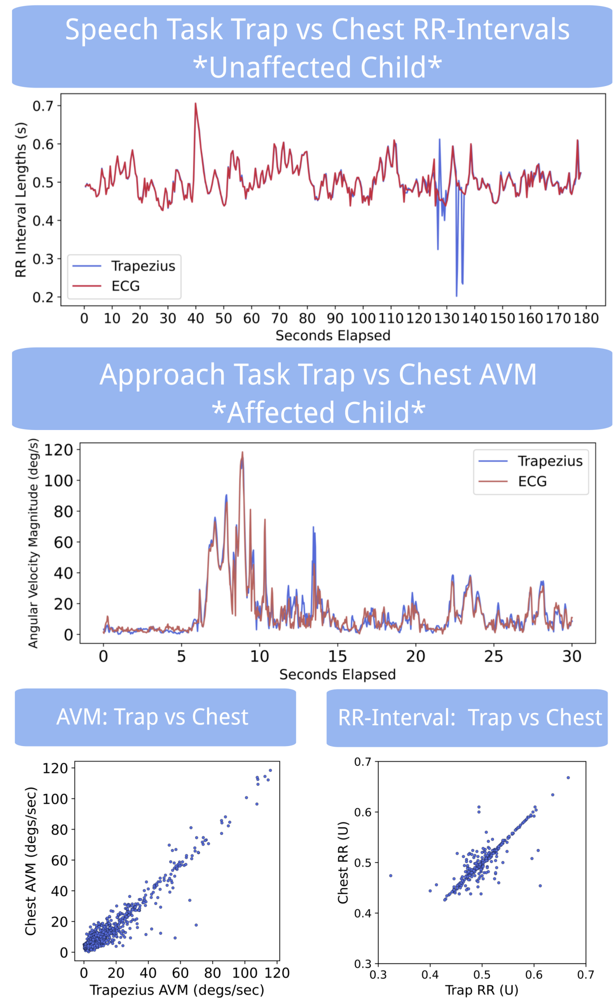

Figure 3. Signal redundancy for movement and R-R interval between trap and chest BioStamps. R-R interval (timeseries top, scatter plot bottom right) and movement (timeseries middle, scatter plot bottom left) are highly correlated between the two devices during the Speech and Approach tasks.

Herein, we identified candidate features of anxiety during the Speech Task. Our results suggest that increased movement in response to startling stimuli (buzzer sounds) may be a significant discerning feature. This is supported by our previous finding that increased vocal pitch in response to buzzers was a discerning feature of internalizing disorders [9]. Similarly, previous studies have identified eyeblink magnitude in response to unexpected stimuli was a discerning feature of phobia-related anxiety disorders in adults [21]. Our results also suggest that R-R interval reactivity during the Speech Task may be a significant discerning feature. This is supported by previous work demonstrating heightened cortisol reactivity in adults with internalizing disorders [23]. Collectively, these results demonstrate that the proposed instrumentation is feasible for use in children and provides physiological measures that may prove useful for building machine learning-based methods for detecting internalizing disorders in young children.

We identified a possible point of redundancy in our wearable sensor array. The strong signal correlations of the chest and trap sensors indicate the potential for removing the chest sensor and considering just the trap. This would enable simultaneous collection of heart rate, heart rate variability, trap EMG, and torso movement data using a single device. This reduction may be especially relevant for our very young participants, as several of them have noted that the sensor adhesive feels too strong for their sensitive skin.

V. Conclusion

These findings will inform our research in this on-going study geared toward developing a digital tool for detecting childhood internalizing disorders. Overall, startle movement reactivity and R-R interval reactivity during the Speech Task are identified as candidate features for future modeling efforts. Decreasing the number of wearable sensors now appears plausible for more feasible deployment. These findings further our work toward better identification and tracking of internalizing mental health disorders across childhood aligning with initiatives by the NIMH Strategic Plan for Research 2021, the American Academy of Pediatrics and US Preventative Services Task Force.

Acknowledgments

This work would not be possible without the support of the KID Study undergraduate researchers: Taylor Mullen, Nadia Huber, Jocelyn Botelho, Lucy Welch, Antranig Douglas, and Haley Meskin.

References

[1] H. L. Egger and A. Angold, “Common emotional and behavioral disorders in preschool children: presentation, nosology, and epidemiology,” *J Child Psychol Psychiatry*, vol. 47, no. 3–4, pp. 313–337, Apr. 2006, doi: 10.1111/j.1469-7610.2006.01618.x.

[2] M. Tandon, E. Cardeli, and J. Luby, “Internalizing Disorders in Early Childhood: A Review of Depressive and Anxiety Disorders,” *Child and Adolescent Psychiatric Clinics of North America*, vol. 18, no. 3, pp. 593–610, Jul. 2009, doi: 10.1016/j.chc.2009.03.004.

[3] N. R. Towe-Goodman, L. Franz, W. Copeland, A. Angold, and H. Egger, “Perceived Family Impact of Preschool Anxiety Disorders,” *J Am Acad Child Adolesc Psychiatry*, vol. 53, no. 4, pp. 437–446, Apr. 2014, doi: 10.1016/j.jaac.2013.12.017.

[4] T. E. Chansky and P. C. Kendall, “Social expectancies and self-perceptions in anxiety-disordered children,” *J Anxiety Disord*, vol. 11, no. 4, pp. 347–363, Aug. 1997.

[5] D. J. Kolko and A. E. Kazdin, “Emotional/behavioral problems in clinic and nonclinic children: Correspondence among child, parent and teacher reports,” *Journal of Child Psychology and Psychiatry*, vol. 34, no. 6, pp. 991–1006, 1993, doi: 10.1111/j.1469-7610.1993.tb01103.x.

[6] A. G. Renouf and M. Kovacs, “Concordance between mothers’ reports and children’s self-reports of depressive symptoms: A longitudinal study,” *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 33, no. 2, pp. 208–216, Feb. 1994, doi: 10.1097/00004583-199402000-00008.

[7] C. Weitzman, L. Wegner, and C. on P. A. of C. and F. H. the SECTION ON DEVELOPMENTAL AND BEHAVIORAL PEDIATRICS, “Promoting Optimal Development: Screening for Behavioral and Emotional Problems,” *Pediatrics*, vol. 135, no. 2, pp. 384–395, Feb. 2015, doi: 10.1542/peds.2014-3716.

[8] R. S. McGinnis *et al.*, “Rapid detection of internalizing diagnosis in young children enabled by wearable sensors and machine learning,” *PLOS ONE*, vol. 14, no. 1, p. e0210267, Jan. 2019, doi: 10.1371/journal.pone.0210267.

[9] E. W. McGinnis *et al.*, “Giving Voice to Vulnerable Children: Machine Learning Analysis of Speech Detects Anxiety and Depression in Early Childhood,” *IEEE Journal of Biomedical and Health Informatics*, pp. 1–1, 2019, doi: 10.1109/JBHI.2019.2913590.

[10] R. S. McGinnis *et al.*, “Rapid Anxiety and Depression Diagnosis in Young Children Enabled by Wearable Sensors and Machine Learning,” presented at the 2018 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Honolulu, HI, Jul. 2018.

[11] E. W. McGinnis *et al.*, “Wearable sensors detect childhood internalizing disorders during mood induction task,” *PLOS ONE*, vol. 13, no. 4, p. e0195598, Apr. 2018, doi: 10.1371/journal.pone.0195598.

[12] R. S. McGinnis *et al.*, “Wearable Sensors and Machine Learning Diagnose Anxiety and Depression in Young Children,” presented at the 2018 IEEE International Conference on Biomedical and Health Informatics (BHI), Las Vegas, NV, Mar. 2018.

[13] E. McGinnis *et al.*, “Movements indicate threat response phases in children at-risk for anxiety,” *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 5, pp. 1460–1465, 2017, doi: 10.1109/JBHI.2016.2603159.

[14] B. N. Cuthbert and T. R. Insel, “Toward the future of psychiatric diagnosis: the seven pillars of RDoC,” *BMC Med*, vol. 11, p. 126, 2013, doi: 10.1186/1741-7015-11-126.

[15] J. S. Silk, D. S. Shaw, E. E. Forbes, T. L. Lane, and M. Kovacs, “Maternal Depression and Child Internalizing: The Moderating Role of Child Emotion Regulation,” *Journal of Clinical Child & Adolescent Psychology*, vol. 35, no. 1, pp. 116–126, Feb. 2006, doi: 10.1207/s15374424jccp3501_10.

[16] J. S. Moser, C. E. Durbin, C. J. Patrick, and N. B. Schmidt, “Combining Neural and Behavioral Indicators in the Assessment of Internalizing Psychopathology in Children and Adolescents,” *J Clin Child Adolesc Psychol*, Jan. 2014, doi: 10.1080/15374416.2013.865191.

[17] J. R. Gagne, C. A. Van Hulle, N. Aksan, M. J. Essex, and H. H. Goldsmith, “Deriving childhood temperament measures from emotion-eliciting behavioral episodes: scale construction and initial validation,” *Psychol Assess*, vol. 23, no. 2, pp. 337–353, Jun. 2011, doi: 10.1037/a0021746.

[18] M. Gilliom, D. S. Shaw, J. E. Beck, M. A. Schonberg, and J. L. Lukon, “Anger regulation in disadvantaged preschool boys: strategies, antecedents, and the development of self-control.,” *Developmental psychology*, vol. 38, no. 2, p. 222, 2002.

[19] J. Pan and W. J. Tompkins, “A Real-Time QRS Detection Algorithm,” *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985, doi: 10.1109/TBME.1985.325532.

[20] S. Chatterjee, R. S. Thakur, R. N. Yadav, L. Gupta, and D. K. Raghuvanshi, “Review of noise removal techniques in ECG signals,” *IET Signal Processing*, vol. 14, no. 9, pp. 569–590, 2020, doi: 10.1049/iet-spr.2020.0104.

[21] L. M. McTeague and P. J. Lang, “The Anxiety Spectrum and the Reflex Physiology of Defense: From Circumscribed Fear to Broad Distress,” *Depression and Anxiety*, vol. 29, no. 4, pp. 264–281, 2012, doi: 10.1002/da.21891.

[22] A. Buske-Kirschbaum, S. Jobst, A. Wustmans, C. Kirschbaum, W. Rauh, and D. Hellhammer, “Attenuated free cortisol response to psychosocial stress in children with atopic dermatitis,” *Psychosom Med*, vol. 59, no. 4, pp. 419–426, Aug. 1997.

[23] K. Roelofs, J. van Peer, E. Berretty, P. de Jong, P. Spinhoven, and B. M. Elzinga, “Hypothalamus–Pituitary–Adrenal Axis Hyperresponsiveness Is Associated with Increased Social Avoidance Behavior in Social Phobia,” *Biological Psychiatry*, vol. 65, no. 4, pp. 336–343, Feb. 2009, doi: 10.1016/j.biopsych.2008.08.022.

[24] M. R. Gunnar, N. M. Talge, and A. Herrera, “Stressor paradigms in developmental studies: What does and does not work to produce mean increases in salivary cortisol,” *Psychoneuroendocrinology*, vol. 34, no. 7, pp. 953–967, Aug. 2009, doi: 10.1016/j.psyneuen.2009.02.010.