Multimodal Markers of Transdiagnostic Childhood Mental Health Impairment

Josh Cherian
Center for Remote Health Monitoring
Wake Forest School of Medicine
Winston-Salem, NC, United States
0000-0002-7749-2109

Bryn Loftness
M-Sense Research Group
University of Vermont
Burlington, VT, United States
0000-0003-4597-0783

Jenna Cohen
M-Sense Research Group
University of Vermont
Burlington, VT, United States
0009-0004-7189-9663

Julia Halvorson-Phelan
Medical Center Department of Psychiatry
University of Vermont
Burlington, VT, United States
0000-0003-1654-9228

Ellen W. McGinnis
Center for Remote Health Monitoring
Wake Forest School of Medicine
Winston-Salem, NC, United States
0000-0001-8566-2289

Ryan S. McGinnis
Center for Remote Health Monitoring
Wake Forest School of Medicine
Winston-Salem, NC, United States
0000-0001-8396-6967

2024 IEEE 20th International Conference on Body Sensor Networks (BSN) | 979-8-3315-3014-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/BSN63547.2024.10780528

***Abstract*—Childhood mental health problems are impairing, predictive of health problems later in life, and becoming increasingly prevalent. A critical first step toward addressing this growing crisis is facilitating more widespread screening; however, gold-standard assessments remain subjective, time-consuming, largely unable to capture subthreshold conditions and comorbidity, and limited by access to the clinical experts needed to provide and interpret the results. In response, researchers have developed digital phenotype screening tools that capture and utilize objective physiological and behavioral measures to augment traditional mental health screening. However, the efficacy of these tools has traditionally been evaluated on their ability to predict mental health diagnoses. In contrast, in this work and in line with the Research Domain Criteria (RDoC) framework, we explore the utility of these objective measures for quantifying transdiagnostic severity of impairment across a range of widely used clinical measurements. Using canonical correlation analysis, we find that linear combinations of movement and audio features extracted from smartphone sensor data collected during short ($<$ 7 minutes) objective assessments are significantly correlated with a range of widely used clinical measurements. These findings suggest that easy-to-collect objective physiological and behavioral measures are indicative of severity across a range of psychopathologies.**

***Index Terms*—Digital health, digital biomarkers, pediatric mental health, mobile health**

I. Introduction

The incidence of emotional and mental health disorders in children has increased in recent decades [1]. These conditions impair functioning, relationships, and development, and without treatment have been linked to the development of a range of health problems in adolescence and adulthood [2], [3]. Indeed studies have shown that detecting mental health conditions and intervening appropriately early has a significant positive impact on later health outcomes and life quality [4]. Nevertheless, mental health problems among children remain frequently undetected and only half of the nearly 8 million children in the U.S. with a mental health disorder receive treatment [1].

In response to this growing epidemic, guidance from the American Academy of Pediatrics and the U.S. Preventative Services Task Force has recommended that every child receive mental health screening [5], [6]. Current gold-standard screening methods used by clinicians are predominantly categorical, relying on techniques including caregiver-report questionnaires and structured interviews to ascertain whether a child meets the criteria for a specific disorder. This approach to diagnosing mental health conditions is subjective [7], [8] and time-consuming [9]. Furthermore, by treating mental health conditions as binary states, these screening methods are largely unable to diagnose comorbid and subthreshold disorders, which have high incidence rates and are impairing and developmentally debilitating [10], [11]. Fundamentally, these shortcomings indicate that a significant number of children suffer from mental health conditions that are not diagnosed appropriately, and thus that they do not receive the care they need [12].

To address these weaknesses, there has been a concerted effort among researchers to develop accurate, objective, and widely accessible tools for identifying mental health conditions in young children. The Research Domain Criteria (RDoC) matrix has been a particularly helpful framework in this endeavor, associating variations in behavioral and neurobiological processes with mental health. In contrast to existing clinical measures, which are designed to identify the presence of specific disorders, RDoC is designed to be both transdiagnostic and dimensional, seeking to understand common pathways and processes that contribute to various psychopathologies across multiple levels of analysis (e.g., genes, cells, behavior, self-reports) [13]. In the years since the introduction of this framework, researchers have demonstrated the merit of this design, showing that objective measures captured during mood induction tasks, which are short, monitored exercises designed to elicit specific behavioral responses (e.g., asking participants to tell a story within a short period of time to elicit anxiety/fear), consistently correspond to the mechanisms underlying a broad range of psychopathologies [14].

In our prior work, we demonstrated that digital phenotypes captured from wearable sensors and smartphones during short mood induction tasks are associated with specific mental health disorders and have shown significant promise in ameliorating the issues with existing gold-standard surveys [15]–[19]. In this work, we take this a step further, demonstrating that these objective measures are significantly correlated with a range of behavioral and clinical screening questionnaires. This suggests that these markers are indicative of transdiagnostic severity of mental health impairment.

II. Methods

A. Clinical Measures

In this work, we explore the relationship between objective measures captured during mood induction tasks and an array of widely used clinical measures. In terms of clinical measures, we look at the following caregiver-report surveys: the Behavioral Inhibition Questionnaire (BIQ) [20] and the Child Behavior Questionnaire Very Short form (CBQ) [21] which measure aspects of child temperament, and the Child Behavior Checklist (CBCL) [22] and the SPENCE Children's Anxiety Scale [23] which measure emotional and behavioral problems. Additionally, we look at the Kiddie-Schedule for Affective Disorders and Schizophrenia (KSADS) [24] which is a semi-structured diagnostic interview that is administered to caregivers and asks about their child's behavior over the last 3 months and measures mental health symptoms. We look at the following subscales within each of those scales: BIQ: social novelty (adults, peers, performance, total) and situational novelty (separation, new situation, physical challenge, total); CBCL: internalizing (anxious/depressed, withdrawn, somatic complaints) and externalizing (inattention, aggression); CBQ: surgency, negative affect, and effortful control; KSADS: depression, anxiety, OCD, ADHD, and ODD; SPENCE: separation anxiety, social phobia, obsessive-compulsive, panic, physical injury fears, and generalized anxiety.

B. ChAMP Assessment Battery

The Childhood Assessment and Management of digital Phenotypes (ChAMP) Assessment Battery (shown in Fig. 1) consists of three mood induction tasks designed to replicate the negative and positive valence constructs of the RDoC framework. The Approach Task, designed to elicit fear and anxiety, has the research administrator lead the child toward the back of the study room where an object is covered with a blanket until the child is within one foot of the covered object, at which point the administrator removes the blanket indicating that nothing was underneath. The Speech Task, designed to elicit anxiety, has the child tell a story for three minutes; a buzzer sounds at the halfway point and with 30 seconds remaining. The Bubbles Task, designed to understand initial and sustained reward responsiveness, has the child brought into a room where an administrator turns on a bubble machine for three minutes. There is a three-minute rest period in between each of the tasks.

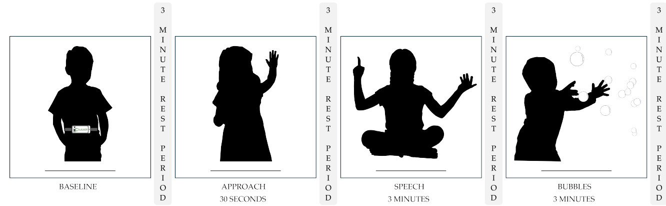

Fig. 1. ChAMP battery protocol consisting of the Approach, Bubbles, and Speech tasks. Tasks are separated by 3-minute rest periods. Data is collected from the ChAMP app running on a smartphone worn on the small of the back. The battery is preceded by a 5-second Baseline period that provides data for sensor calibration.

C. Data Overview

We collected sensor data from 104 children 4-8 years old (M: 6.79 years old; SD: 1.21, 38.5% female) as they went through the ChAMP Assessment battery. For this work, we used movement features extracted from accelerometer and gyroscope data captured during the Approach and Bubbles task and audio features extracted from microphone data captured during the Speech task. We used data from 92 children: data from children who did not have sensor data for one or more of the tasks were excluded. We extracted a set of 406 features (77 approach, 77 bubbles, 252 speech) from captured sensor data. These data and features have been described previously [25].

The feature set was reduced using PCA such that 75% of the variance was explained for each task, resulting in 4 components for the Approach task, 11 for Bubbles, and 13 for Speech. We used canonical correlation analysis (CCA) to understand the relationship between these components and each of the various clinical measures.

III. Results

Features extracted from the performance of each of the tasks were significantly correlated with one or more clinical measures; however, movement features extracted from the performance of the approach task were generally significantly correlated with fewer clinical measures than the bubbles or speech tasks. Notably, audio features were significantly correlated with all clinical measures with the exception of SPENCE Social Phobia ($p = 0.05$). Significant correlations for each clinical measure group and task are shown in Fig. 2.

To understand how these objective measures were associated with clinical measures, we looked at boys vs. girls and younger kids (ages 5-6) vs. older kids (ages 7-8). Fig. 3 shows the percentage of clinical measures that were significantly correlated with features for each clinical measure and task. Features extracted from the performance of all three tasks were significantly correlated with as many or more clinical measures for boys than for girls with the exception of KSADS for the Bubbles and Speech tasks. Similarly, features extracted from the performance of all three tasks were significantly correlated with as many or more clinical measures for older kids than

Fig. 2. Correlations obtained from CCA between motion and audio features and BIQ, CBCL, CBQ, KSADS, and Spence measures for the Approach, Speech, and Bubbles tasks. Non-gray bars had significant correlations ($p < 0.05$).

Fig. 3. Percentage of significant correlations obtained from CCA between motion and audio features and BIQ, CBCL, CBQ, KSADS, and Spence measures for the Approach, Speech, and Bubbles tasks across subgroups.

for younger kids with the exception of SPENCE for the Speech task. Notably, both overall and across all demographic subgroups features extracted from the bubbles and speech task were correlated with at least one clinical measure.

IV. Discussion

Existing techniques for diagnosing mental health conditions in children rely on subjective and categorical assessments that fail to identify the presence of disorders in children who do not meet rigid diagnostic thresholds. This means that a significant number of children with one or more mental health disorders do not get diagnosed or get the treatment they need. We examined the association between audio and movement features extracted during mood induction tasks and clinical measures, finding that in many cases there was a significant correlation. This suggests that this assessment is capturing and measuring many of the mechanisms that underlie a broad range of psychopathologies, and as a result can serve as a short, objective, transdiagnostic approach to understanding an individual's mental health.

The three mood induction tasks comprising the ChAMP battery were designed to objectively measure positive and negative valence in children, and our finding that these tasks are correlated with multiple clinical measures is corroborated by a number of studies that have shown that subdomains of these valences are linked to both internalizing and externalizing disorders [26], [27]. We did find that the objective measures captured during the Approach task are significantly correlated with fewer clinical measures than in the case of the Bubbles and Speech tasks. Specifically, we found significant correlations with clinical measures that looked at how children respond to new situations, potential physical challenges, and their extroversion. While the Speech task was also designed to measure negative valence, it primarily induces a social aspect of fear and anxiety in contrast to the more 'primal/defensive' fear elicited by the Approach task. It is worth highlighting here that our results indicate that audio features captured during the performance of the speech task were significantly correlated with the broadest range of clinical measures, likely because it captures this social aspect of fear and anxiety but incorporates the startle response of the Approach task with the inclusion of the buzzer at several intervals without the participant's prior knowledge. The movement features captured during the performance of the Bubbles task were significantly correlated with most clinical measures, indicating that expressions of positive valence are also highly indicative of a variety of psychopathologies. Future work should explore whether these tasks are individually sufficient to objectively and robustly measure the mechanisms underlying mental health disorders (specifically the Speech and Bubbles tasks given their association with most clinical measures), or if the combination of features from multiple tasks produces more descriptive and robust objective measures. Looking across subgroups, we did find that regardless of the task, objective measures were significantly correlated with fewer clinical measures for younger kids and girls; future work should investigate the associations of these measures with clinical measures within larger samples of these demographic subgroups.

V. Conclusion

Accurate and objective screening of mental health disorders in children is critical to ensuring that they receive the help they need. Objective measures captured during mood induction tasks have been shown to be predictive of specific diagnoses; however, their utility as broader assessments of transdiagnostic severity remains uninvestigated. In this work we demonstrate that movement and vocal biomarkers collected during short mood induction tasks are significantly correlated with a range of clinical measures, suggesting that these features and tasks are capturing mechanisms underlying internalizing and externalizing disorders in children.

References

[1] D. G. Whitney and M. D. Peterson, "US national and state-level prevalence of mental health disorders and disparities of mental health care use in children," *JAMA pediatrics*, vol. 173, no. 4, pp. 389–391, 2019.

[2] A. Bittner, H. L. Egger, A. Erkanli, E. Jane Costello, D. L. Foley, and A. Angold, "What do childhood anxiety disorders predict?" *Journal of Child Psychology and psychiatry*, vol. 48, no. 12, pp. 1174–1183, 2007.

[3] R. M. Rapee, C. A. Schniering, and J. L. Hudson, "Anxiety disorders during childhood and adolescence: Origins and treatment," *Annual review of clinical psychology*, vol. 5, pp. 311–341, 2009.

[4] S. E. Fox, P. Levitt, and C. A. Nelson III, "How the timing and quality of early experiences influence the development of brain architecture," *Child development*, vol. 81, no. 1, pp. 28–40, 2010.

[5] M. Viswanathan, I. F. Wallace, J. C. Middleton, S. M. Kennedy, J. McKeeman, K. Hudson, C. Rains, E. B. Vander Schaaf, and L. Kahwati, "Screening for anxiety in children and adolescents: evidence report and systematic review for the us preventive services task force," *JAMA*, vol. 328, no. 14, pp. 1445–1455, 2022.

[6] J. M. Foy, A. A. of Pediatrics Task Force on Mental Health *et al.*, "Enhancing pediatric mental health care: report from the american academy of pediatrics task force on mental health. introduction," *Pediatrics*, vol. 125, pp. S69–S74, 2010.

[7] 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, 1994.

[8] E. W. McGinnis, W. Copeland, L. Shanahan, and H. L. Egger, "Parental perception of mental health needs in young children," *Child and adolescent mental health*, vol. 27, no. 4, pp. 328–334, 2022.

[9] A. L. Olson, K. J. Kelleher, K. J. Kemper, B. S. Zuckerman, C. S. Hammond, and A. J. Dietrich, "Primary care pediatricians' roles and perceived responsibilities in the identification and management of depression in children and adolescents," *Ambulatory Pediatrics*, vol. 1, no. 2, pp. 91–98, 2001.

[10] P. Fusar-Poli, M. Solmi, N. Brondino, C. Davies, C. Chae, P. Politi, S. Borgwardt, S. M. Lawrie, J. Parnas, and P. McGuire, "Transdiagnostic psychiatry: a systematic review," *World Psychiatry*, vol. 18, no. 2, pp. 192–207, 2019.

[11] P. M. Lewinsohn, S. A. Shankman, J. M. Gau, and D. N. Klein, "The prevalence and co-morbidity of subthreshold psychiatric conditions," *Psychological medicine*, vol. 34, no. 4, pp. 613–622, 2004.

[12] J. L. Luby, A. K. Heffelfinger, C. Mrakotsky, K. M. Brown, M. J. Hessler, J. M. Wallis, and E. L. Spitznagel, "The clinical picture of depression in preschool children," *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 42, no. 3, pp. 340–348, 2003.

[13] B. N. Cuthbert and T. R. Insel, "Toward the future of psychiatric diagnosis: the seven pillars of RDoC," *BMC medicine*, vol. 11, pp. 1–8, 2013.

[14] H. B. Simpson, "The RDoC project: a new paradigm for investigating the pathophysiology of anxiety," pp. 251–252, 2012.

[15] E. W. McGinnis *et al.*, "Wearable sensors detect childhood internalizing disorders during mood induction task," *Plos one*, vol. 13, no. 4, p. e0195598, 2018.

[16] E. W. McGinnis, S. P. Anderau, J. Hruschak, R. D. Gurchiek, N. L. Lopez-Duran, K. Fitzgerald, K. L. Rosenblum, M. Muzik, and R. S. McGinnis, "Giving voice to vulnerable children: machine learning analysis of speech detects anxiety and depression in early childhood," *IEEE journal of biomedical and health informatics*, vol. 23, no. 6, pp. 2294–2301, 2019.

[17] R. S. McGinnis, E. W. McGinnis, J. Hruschak, N. L. Lopez-Duran, K. Fitzgerald, K. L. Rosenblum, and M. Muzik, "Rapid detection of internalizing diagnosis in young children enabled by wearable sensors and machine learning," *PloS one*, vol. 14, no. 1, p. e0210267, 2019.

[18] E. W. McGinnis, J. Scism, J. Hruschak, M. Muzik, K. L. Rosenblum, K. Fitzgerald, W. Copeland, and R. S. McGinnis, "Digital phenotype for childhood internalizing disorders: less positive play and promise for a brief assessment battery," *IEEE journal of biomedical and health informatics*, vol. 25, no. 8, pp. 3176–3184, 2021.

[19] A. O'Leary, T. Lahey, J. Lovato, B. Loftness, A. Douglas, J. Skelton, J. G. Cohen, W. E. Copeland, R. S. McGinnis, and E. W. McGinnis, "Using wearable digital devices to screen children for mental health conditions: Ethical promises and challenges," *Sensors*, vol. 24, no. 10, p. 3214, 2024.

[20] G. Bishop, S. H. Spence, and C. McDonald, "Can parents and teachers provide a reliable and valid report of behavioral inhibition?" *Child development*, vol. 74, no. 6, pp. 1899–1917, 2003.

[21] M. K. Rothbart, S. A. Ahadi, K. L. Hershey, and P. Fisher, "Investigations of temperament at three to seven years: The children's behavior questionnaire," *Child development*, vol. 72, no. 5, pp. 1394–1408, 2001.

[22] T. M. Achenbach, L. Dumenci, and L. A. Rescorla, "Ratings of relations between dsm-iv diagnostic categories and items of the cbcl/6-18, trf, and ysr," *Burlington, VT: University of Vermont*, pp. 1–9, 2001.

[23] S. H. Spence, "Spence children's anxiety scale," *Journal of Anxiety Disorders*, 1997.

[24] J. Kaufman, B. Birmaher, D. Brent, U. Rao, C. Flynn, P. Moreci, D. Williamson, and N. Ryan, "Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (k-sads-pl): initial reliability and validity data," *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 36, no. 7, pp. 980–988, 1997.

[25] B. C. Loftness *et al.*, "The ChAMP app: A scalable mhealth technology for detecting digital phenotypes of early childhood mental health," *IEEE Journal of Biomedical and Health Informatics*, 2023.

[26] T. M. Olino, D. N. Klein, M. W. Dyson, S. A. Rose, and C. E. Durbin, "Temperamental emotionality in preschool-aged children and depressive disorders in parents: associations in a large community sample." *Journal of abnormal psychology*, vol. 119, no. 3, p. 468, 2010.

[27] 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.