Multi-stakeholder Perspectives on Mental Health Screening Tools for Children

Manasa Kalanadhabhatta
University of Massachusetts Amherst
Amherst, MA, USA
manasak@cs.umass.edu

Adrelys Mateo Santana
University of Southern California
Los Angeles, CA, USA
amateosa@usc.edu

Lynnea Mayorga
University of Massachusetts Amherst
Amherst, MA, USA
lmayorga@umass.edu

Tauhidur Rahman
University of California San Diego
La Jolla, CA, USA
trahman@ucsd.edu

Deepak Ganesan
University of Massachusetts Amherst
Amherst, MA, USA
dganesan@cs.umass.edu

Adam Grabell
University of Massachusetts Amherst
Amherst, MA, USA
agrabell@umass.edu

ABSTRACT

Pediatric mental health is a growing concern around the world, affecting children's social-emotional development and increasing the risk of poor behavioral outcomes later in life. However, obtaining a behavioral diagnosis in early childhood is challenging due to lack of access to resources, low parental mental health literacy, and children's dependence on several stakeholders to coordinate care for them. While app-based, at-home screening tools could offer a scalable and convenient diagnostic solution for families, stakeholder perspectives on their utility and usability remain to be examined. This work reports on a survey of child mental health practitioners and interviews with parents to illustrate existing barriers to care that stakeholders encounter, the perceived benefits of app-based screening tools in meeting their needs, and the challenges in scaling up these tools. We identify where stakeholders agree or disagree, delineate key design tensions, and provide recommendations for the development of future screening technologies.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; *Ubiquitous and mobile computing*; • **Applied computing** → **Health informatics**.

KEYWORDS

mental health, children, parenting, caregiving, digital health

1 INTRODUCTION

Over 7 million children in the United States, or 16.5% of the population under 18 years of age, are impacted by mental health disorders [90]. Poor mental health in childhood is associated with a decline in quality of life [73], impaired social-emotional development [18, 24], and poor learning outcomes [61]. Childhood psychopathology also predicts persistent risk of mental disorders into adulthood and across the lifespan [69], the subsequent onset of substance use disorders [47], and a negative impact on employment status and income as an adult [11]. Additionally, pediatric mental health exerts an enormous economic burden on society [11, 25].

While this has led to growing interest in diagnosing and treating childhood mental disorders, over half of the children with mental health issues fail to receive treatment [77, 90]. The mental health treatment gap in pediatric populations stems from a range of systemic issues, including an under-resourced mental health system [51], socioeconomic and racial disparities [21, 55], the stigma associated with mental disorders [5], and logistical or financial burdens of accessing care [17]. Children specifically also cannot seek help for themselves and require an adult primary caregiver to identify potential problems and further have the institutional knowledge and resources to address them [58]. However, parents and/or family caregivers may not be able to distinguish between normative behavior and signs of non-normative development in young children, creating a "when to worry" problem for families [86]. Evidence-based clinical assessments for young children are also difficult and time-consuming to administer and have relatively poor diagnostic accuracy and specificity [37, 57].

Digital mental health tools have been recognized as alternatives or supplements to the existing mental health care landscape. A range of technologies have been developed to support families track and manage children's development and behavioral symptoms (e.g., [43, 46, 48, 59, 64]). Digital interventions have been proposed for depression, anxiety, attention-deficit/hyperactivity disorder (ADHD), autism spectrum disorder, etc. in adolescents and young adults [35]. Researchers have highlighted the importance of co-designing digital health interventions with users to increase their uptake and continued use, user satisfaction, and the likelihood of success [14, 28, 82]. This has led to several design explorations and user-centered development exercises in the human-computer interaction (HCI) and computer-supported cooperative work (CSCW) communities around child mental health (e.g., [14, 52, 71, 83]).

However, the design space of mental health *screening* technologies for children, especially toddlers or preschoolers, has been relatively under-explored. Prior research has focused on understanding technology use practices [29] and decision-making experiences [50] among parents of infants and toddlers. Concerning physical health, the development of caregivers' experiential knowledge [75] and tensions around parenting children with chronic medical needs [74] have been explored previously. Researchers have further looked into how technology can support new parents [30, 31, 48, 85], but there remains a lack of understanding regarding how sociotechnical systems can support caregivers of children with behavioral concerns identify and seek help for these concerns. Similarly, although there has been prior work exploring tools for parent-clinician communication in the context of children's health (eg., [43, 64]), it is unclear what data child mental health practitioners would want to collect using at-home sensing technologies.

In this work, we attempt to fill these gaps by examining the perspectives of two key stakeholders – child mental health professionals/clinicians and parents/family caregivers – toward using mental health screening or diagnostic applications for young children[1]. Specifically, we focus on investigating the perceived utility and drawbacks of **app-based, at-home screening tools that utilize behavioral data to predict mental health outcomes in preschool-aged children**. Our motivation to concentrate on this specific class of screening technologies stems from prior work demonstrating the increasing use of mobile applications to support parenting, especially in the early years of a child's life [85]. Previous HCI research has also made significant advances in developing app-based screening tools for young children using a variety of sensing modalities (e.g., [36, 43, 44, 48]), but comparatively little is known about stakeholders' perspectives towards such tools.

Within the space of app-based screening tools, our work attempts to answer the following research question: **What are the unmet needs and gaps in current caregiving practices of parents and clinicians, and to what extent can these be addressed by at-home screening technologies?** Recognizing the needs of both stakeholder groups is imperative for researchers and technology developers to create tools that are beneficial and amenable to users and can be seamlessly integrated with existing diagnostic practices. We therefore also investigate differences in opinion among parents and clinicians concerning child mental health screening tools.

We analyze free-form survey responses from sixty mental health professionals collected by the authors of a prior work that introduced an app-based screening tool for young children [44]. Our study is, to the best of our knowledge, the first to undertake a systematic qualitative inquiry using this dataset to derive thematic findings. We complement this with original data from interviews with 26 parents of preschool-aged children to gain insight into their lived experience identifying and managing behavioral issues.

Thematic analysis of survey and interview data reveals a lack of scaffolding to support parental help-seeking, a dearth of resources for tracking emotional development, and limitations in data availability that impact pediatric mental health screening. Stakeholders collectively outlined various benefits of app-based screening tools, including increased access to screening, the potential to augment existing practices, making diagnosis more frictionless for children, and alleviating parental pressure. At the same time, they identified challenges in deploying such tools, including integrating them with traditional services, building trust, and minimizing potential harm. Our work also delineates how clinicians and parents have different values and sensitivities concerning the utility and drawbacks of home-based screening tools. We further analyze stakeholders' perspectives through Tatar's design tensions framework [81] to identify conflicting requirements that must be balanced against each other to build highly scalable and usable mental health screening tools. We offer concrete design recommendations for developers of screening technologies to achieve this goal.

2 RELATED WORK

2.1 Technologies for Development Tracking in Early Childhood

There has been long and sustained interest in the HCI community to develop tools that support parents in tracking their child's development in the first few years of life. For example, Baby Steps encouraged parents to frequently record their child's developmental progress and share it with their pediatrician [48]. The system was further extended to a Twitter-based service [78] as well as a text message-based developmental screening tool [79]. babyTRACKS allowed parents to record short textual descriptions of developmental milestones and receive expert-curated, crowd-based percentiles to compare their child's development relative to others [12]. SpecialTime analyzed parent-child dialogues to provide parents with real-time feedback on dyadic interaction patterns while the family underwent parent-child interaction therapy [36]. Research has additionally looked into collaboration opportunities between parents and healthcare providers for identifying relevant data to track [83].

Other tools have focused on communication and information sharing among various stakeholders involved in caring for young children. For example, CRAFT allowed parents of children with developmental disabilities to record videos of problem instances that were later reviewed by behavior analysts [64]. GeniAuti supported parents of autistic children in recording challenging behaviors based on a clinical data collection form, which experts could then use to make recommendations [43]. In light of the significant progress made by the above technologies toward identifying and communicating mental health concerns, our work aims to systematically investigate multi-stakeholder perspectives on the perceived utility and impact of such screening tools in meeting their current needs.

2.2 Play- and Game-based Mental Health Screening Tools

Prior work has further attempted to automatically detect developmental delays or screen for behavioral disorders among young children. Boccanfuso et al. used play patterns and affective responses of

---

[1]Throughout this work, we use the terms "mental health professionals", "clinicians", and "practitioners" to refer to individuals who are trained and licensed to provide mental health care to children and families as part of the behavioral health care system (e.g., psychologists, psychiatrists, licensed social workers, etc.). We use "parents" or "caregivers" to identify members of the child's family who undertake primary caregiving responsibilities, including biological/adoptive parents and/or grandparents. Both groups are together referred to as "stakeholders", as they play distinctive yet important roles in supporting the child's mental health. In the context of the studies described later, both groups are also collectively referred to as "participants".

children while interacting with an emotional robot to differentiate between children with autism spectrum disorder and those developing typically [15]. Building blocks embedded with inertial sensors have been used to sense play styles and predict behavior problems in young children [89]. Researchers have also used accelerometers placed on the wrist and ankle to classify children with ADHD vs. healthy controls during a computerized continuous performance task [32] or over the course of a regular school day [63].

Among game-based screening tools for mental health issues in children, EarlyScreen used video-based behavioral markers to predict emotion regulation-related disorders [44]. Jiang et al. used a combination of interactive devices and wearables to diagnose symptoms of ADHD in 7- to 13-year-old children as they completed gamified tasks on a large touch-screen system [42]. Other researchers have used gameplay data to detect ADHD, anxiety, and depressive disorders [34, 72]. Song et al. developed a narrative game to test various dimensions of cognitive control among children and adolescents [76]. Smartphone-based games have also been developed for in-the-wild detection of autism spectrum disorders in children [22, 84]. In addition to games that have been specifically developed for mental health assessment, Mandryk et al. argue that in-the-wild data collected while playing commercial, off-the-shelf games that are primarily designed for entertainment purposes can be valuable digital biomarkers for mental health [54]. Prior work has also explored associations between older players' mental health and their choice of games and genres [67] as well as in-game behaviors [13].

2.3 Designing Screening Tools to Support Caregivers

While there is limited research on caregivers' use of, and expectations from, technologies for *behavioral screening*, we can draw insight into caregiver motivations from prior work exploring these themes in the context of tools for infant tracking. Marx and Steeves observed that parents are often coerced into tracking or "surveilling" their children using technology in order to keep them safe while promoting parental convenience [56]. This is brought on by both intrinsic and external factors – Lupton notes that mothers are inherently acutely aware of their parenting responsibilities and additionally face societal pressure to "conform to the ideal of the 'good mother' " [53]. Infant tracking technologies can indeed support parents in fulfilling these responsibilities and help alleviate pressure – prior research shows how parents use tools both to confirm hypotheses about their child's development that arise from their own lived experience and to minimize relying solely on their intuition to identify potential concerns [88].

However, technologies that are designed to help parents overcome their anxiety have also been shown to exacerbate it instead, often by providing continuous and unprecedented access to health-related data [45, 65, 88]. Over-reliance on technology may also lead to a false sense of security [88] or complacency and lack of vigilance among parents [10]. It may take away from an embodied parenting experience that relies on knowledge and intuition gained from lived experience [53]. Therefore, technologists have called for tools that support families by improving parents' confidence in fulfilling their new responsibilities, rather than replacing parental

intuition [31]. One such framework is proposed by Kaziunas et al., who advocate for a 'caring-through-data' approach where data from tracking technologies is used to empower users and foster empathy and communication [45].

Most of the above research, however, focuses on supporting parents in tracking physical health and development, ignoring potential concerns about behavioral and social-emotional needs. We aim to fill this gap by examining the perspectives of caregivers toward behavioral screening tools, focusing on app-based, at-home screening technologies that can potentially supplement existing mental health services.

3 METHODS

As described previously, we focus our work on understanding the perspectives of two user groups – mental health practitioners and parents/caregivers – toward app-based behavioral screening tools for young children. Stakeholders' perspectives on screening tools were gathered via a practitioner survey undertaken by prior work and original interviews with parents, which are described in Sections 3.1 and 3.2 respectively.

3.1 Study 1: Analyzing Survey Responses from Child Mental Health Practitioners

For our first study, we qualitatively analyzed survey responses from 60 child clinicians and other mental health practitioners collected by the authors of [44] to examine the perspectives of mental health professionals toward screening apps. While this prior work conducted a quantitative examination of Likert scale responses from clinicians, the present study qualitatively analyzed clinicians' free-form, open-ended comments to derive themes related to the use of digital screening tools, providing a richer picture of how child clinicians in the United States perceive the potential adoption of these tools. We reproduce relevant details about the study methods below.

3.1.1 Participants. Participants were recruited from the mailing list for the Society of Clinical Child and Adolescent Psychology division of the American Psychological Association [9]. While the survey was open to mental health practitioners who worked/had worked with children of any age, we limited our analysis to respondents who have experience working with children under the age of 5 years (since we specifically focus on mental health screening tools for young children in this work). This subset contained 25 clinicians; 22 self-identified as female and 3 as male.

The respondents' age, training, experience, and theoretical outlooks are summarized in Figure 1. Most practitioners held a Doctorate in psychology and had been working as a clinician for over 5 years after earning their highest degree. Their therapy practices were guided primarily by cognitive-behavioral theories, followed by a moderate amount of interpersonal and systems frameworks. Practitioners also reported that they used a developmental lens and applied behavior analysis, and relied heavily on evidence-based assessments to inform evidence-based treatments while also acknowledging their shortcomings.

Clinicians had experience working with children from a wide range of racial and ethnic backgrounds, including Hispanic and/or non-Hispanic white families, African American/Black parents, Asian,
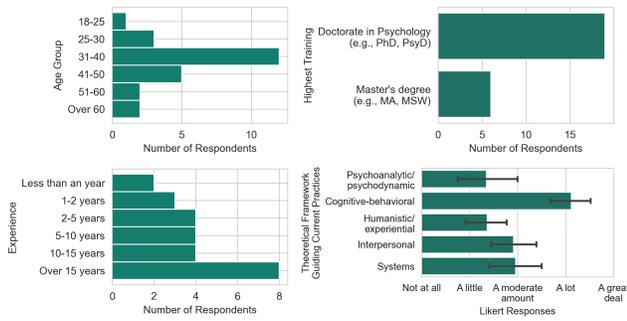
Figure 1: Age, training, experience, and therapy approaches of clinicians who participated in the survey.

South American, and Middle-Eastern immigrant families, and other communities of color. Client families ranged from low or very low socioeconomic status to working-class or middle/upper-middle income. Parents' ages and education levels also varied, and many clinicians saw children from the child welfare or foster care systems, multi-generational homes, and families with international and domestic adoptions.

3.1.2   Survey Context: The EarlyScreen Application. The clinician survey specifically asked participants to comment on EarlyScreen, an at-home behavioral screening app for preschool-aged children developed by the authors of [44]. Briefly, EarlyScreen is a tablet-based app that predicts the child's risk for broad externalizing disorders, including irritability, temper loss, and ADHD. Children play a game where they are introduced to a baker who seeks their help choosing cakes for customers. Customers provide predetermined positive or negative reactions to the child's choices (see Figure 2), with repeated negative feedback meant to induce frustration. The app records children's facial expressions and bodily movements in response to this frustration and uses this data to predict psychopathological risk. For more details on the screening algorithm, we refer readers to the original publication (i.e., [44]).

Survey participants read a textual description of the EarlyScreen application (reproduced in Appendix A) consisting of its purpose and intended use, how it worked, and the predictive performance reported in [44] before answering questions about their perspectives towards the tool. Methodologically, the survey's approach is based on design probes [87] and provocations [2], utilizing an example screening tool to support reflection and elicit reactions from participants.

3.1.3   Survey Procedures. The survey begins by asking about clinicians' backgrounds and current therapy practices without introducing participants to the EarlyScreen probe. Respondents were then invited to describe the gaps and pain points in their current diagnostic practice, with the survey asking *"[w]hat sources of data would you like to have access to as part of your diagnostic process but are often unable to access?"*. Screening technology was still not mentioned at this stage in order to avoid biasing participant responses and instead focus on better understanding their existing needs.

Finally, respondents were presented with a textual description of the EarlyScreen application (see Appendix A). Based on this

Figure 2: The EarlyScreen app [44] consists of a bakery game where children select cakes for each customer. The customers provide predetermined positive or negative feedback. Images courtesy of original authors.

description, clinicians were invited to answer the following questions: (i) *"What do you like most about the EarlyScreen application?"*, (ii) *"What do you like least about the EarlyScreen application?"*, (iii) *"What potential concerns do you have about using or recommending EarlyScreen in addition to your current practices?"*, and (iv) *"What features would you like to see in EarlyScreen that are not described here?"*. All questions were optional and no word limits were imposed.

3.2   Study 2: Interviewing Parents of Preschool-aged Children

In addition to clinician perspectives from the survey described in Section 3.1, we sought to understand the views of another key stakeholder group towards mental health technologies for young children, viz., parents/family caregivers of preschool-aged children. To this end, we conducted semi-structured interviews with parents to gather additional data to answer our research question.

3.2.1   Participants. We recruited 26 caregivers through word-of-mouth and online advertisements within the community as well as playgroups/parent support groups in the area. In order to gather a wide range of parent perspectives, we did not explicitly select parents with existing concerns about their child's behavior or mental health or control for parental mental health literacy. Instead, we sought to interview caregivers from diverse families including single- and dual-income households, rural and urban communities, immigrant and student parents, etc. Table 1 describes the participant demographics in more detail. In most cases, one parent (or another caregiver, e.g., a grandparent) was interviewed one-on-one by an author. In one case, both parents of a child were independently eligible and chose to participate together – they are identified as P4a and P4b. Parents were sometimes interviewed in the presence of the child but the child did not participate in the interview. One parent (P23) was excluded from the analysis due to data recording errors.

3.2.2   The EarlyScreen Application as a Speculative Probe. To maintain consistency with the clinician survey as well as to create a shared understanding for participants with varying levels of familiarity with digital screening tools, we used a video demo of the EarlyScreen application [44] as a speculative probe during our parent interviews. We draw inspiration from both [44] as well as other prior work that uses designed applications and artifacts to evoke users' feedback, thus leveraging the speculative nature of cultural probes [16, 60, 68]. A scenario-based video of EarlyScreen was preferable to using a low- or medium-fidelity prototype of the application as an interactive probe to provoke participants' thoughts for two reasons. First, the direct users of an app like EarlyScreen would

Table 1: Parent Interview Participant Demographics (EA: European-American, H/L: Hispanic or Latino).

Parent ID	Relation to Child	Child's Race	Child's Ethnicity	Household Income	# of Children (Ages)	Concerns about Child's Behavior?
P1	Father	White or EA	Not H/L	$50,000-$74,999	One (3)	No
P2	Mother	Asian	Not H/L	$50,000-$74,999	Two (10, 2)	Yes
P3	Father	Asian	Not H/L	$40,000-$49,999	Two (5, 1.5)	No
P4a/P4b	Mother/Father	Asian	Not H/L	$20,000-$39,999	One (3)	Yes
P5	Father	Asian	Not H/L	$20,000-$39,999	One (2)	No
P6	Father	White or EA	Not H/L	$50,000-$74,999	One (3)	Yes
P7	Father	White or EA	Not H/L	$50,000-$74,999	Three (4, 1.5, 1.5)	Yes
P8	Mother	White or EA	H/L	> $200,000	One (2)	No
P9	Mother	Asian	Not H/L	$40,000-$49,999	Two (8, 6)	Yes
P10	Mother	Biracial	Not H/L	$100,000-$149,999	One (3)	No
P11	Father	White or EA	Not H/L	$50,000-$74,999	Two (5, 3)	Yes
P12	Father	Asian	Not H/L	$20,000-$39,999	One (2)	No
P13	Father	White or EA	Not H/L	$100,000-$149,999	One (4)	Yes
P14	Grandmother	White or EA	H/L	$150,000-$200,000	One (2)	No
P15	Mother	White or EA	Not H/L	$75,000-$99,999	Two (4, 7)	Yes
P16	Mother	White or EA	Not H/L	$100,000-$149,999	Two (3, 7)	Yes
P17	Mother	White or EA	Not H/L	$100,000-$149,999	One (2)	Yes
P18	Grandmother	White or EA	Not H/L	$50,000-$74,999	One (3.5)	Yes
P19	Mother	White or EA	Not H/L	$50,000-$74,999	Four (18, 17, 4, 3)	Yes
P20	Grandmother	White or EA	Not H/L	$20,000-$39,000	One (4)	No
P21	Mother	Biracial	Not H/L	$150,00-$200,000	One (2)	No
P22	Grandmother	White or EA	Not H/L	$20,000-$39,000	Three (6, 4, 0)	Yes
P23	Mother	White or EA	Not H/L	$100,000-$149,000	-	-
P24	Mother	White or EA	Not H/L	$40,000-$49,000	One (2)	Yes
P25	Mother	White or EA	Not H/L	$75,000-$99,999	Two (3, 0)	Yes

be young children, who were not part of this study since we chose to focus on the perspectives of adult stakeholders at this point. Second, video probes could be shared more easily if participants chose to complete the interviews remotely. The one-minute video demo (available as Supplementary Material) described EarlyScreen and how it could be used by parents to help identify whether their child might be at risk for irritability and mood disorders.

3.2.3 Interview Procedures. All parents provided written informed consent before participating. We started the interview by inviting parents to share general concerns about their child's early development and behavior. We asked about caregivers' biggest worries pertaining to their child (*"Do you ever worry about your child? If yes, what is your biggest worry?"*). Parents were not explicitly probed about mental health concerns, rather, they spoke openly about different concerns and family circumstances. If no behavioral concerns were mentioned, parents were asked a single follow-up question (*"Any concerns about their mood or behavior?"*). As listed in Table 1, 15 of the interviewees ultimately expressed behavioral concerns, including but not limited to defiant behavior, temper tantrums, trouble sharing or socializing with peers, separation anxiety, etc.

Parents were then asked about their current help-seeking practices (*"Where do you look for help to ease these worries?"*) and the information gaps that caregivers wanted to fill (*"If you had a crystal ball to learn something about your child and ease your worries, what would it be? What information will you find helpful but just don't*

have access to?"). Screening technologies were not mentioned at this stage and participants were invited to talk about any information they would find helpful.

Following this open-ended conversation, participants were presented with a video demo of EarlyScreen as a speculative probe. Parents were then asked if they would use an app like EarlyScreen (*"If EarlyScreen was widely available and easy to access, would you use it to test your child at home? Why or why not?"*) and/or what concerns they had about it (*"What worries would you have about using such an app?"*). Finally, participants were asked what they would like to see in screening applications in general that would better support their parenting experience (*"What else would you like to see in apps like EarlyScreen that is not described here that would support you as a parent?"*).

All interviews were conducted in English with questions phrased at an elementary-school reading level, and participants were excluded if they could not read, speak, or understand English. Interviews took up to 30 minutes to complete and interviewees received a $15 gift card. Interviews were conducted in person or via video conferencing and the audio was recorded and transcribed for further analysis. The study was approved by the Institutional Review Board at the University of Massachusetts Amherst and permission was sought from playgroup facilitators where applicable.

3.3 Data Analysis

3.3.1 Thematic Analysis. We analyzed survey responses collected by the authors of [44] as well as transcripts from our parent interviews using an inductive approach to identify salient themes, similar to prior work such as [3] and [26]. One author performed rigorous initial (open) coding [20] of the survey and interview data and developed a set of preliminary codes for both sources that were guided by our research question. A codebook was generated using the NVivo qualitative data analysis software [39] and the extracted codes was iteratively refined through discussion among authors. The survey and interview codes were collectively grouped into themes reflecting (i) participants' concerns stemming from their lived experience caring for young children, (ii) their perceived utility of screening tools, and (iii) the challenges of integrating such tools into their existing practices. The rationale for deriving themes using codes from both stakeholder groups was to present a complementary and holistic picture of perspectives toward screening tools (while certain individual themes may still relate to a single stakeholder group). The analysis and resulting themes were reviewed by all authors. We illustrate these themes in Section 4 and list sample codes for each theme in Appendix B.

3.3.2 Differences in Stakeholder Opinions. While a combined thematic analysis of survey and interview codes offers a comprehensive understanding of multi-stakeholder perspectives towards app-based screening tools, clinicians and parents may also sometimes have different or opposing views. We investigated these differences by comparing survey and interview codes to distill the implications of some unique and/or opposing viewpoints. We used NVivo to classify each instance of a code, marking whether it was derived from clinician or parent responses. We compared the frequency of occurrence of codes within each group and discussed the codes and findings among the authors to identify salient topics that are elucidated upon in Section 5.

3.3.3 Design Tensions and Recommendations. We further analyze our findings through a design tensions lens [81] in Section 6.1, identifying conflicting stakeholder values that must be grappled with when designing pediatric mental health screening tools. Briefly, Tatar's design tensions framework conceptualizes the design process as an optimization, or "goal balancing", endeavor in the presence of a "multiplicity of perspectives" or stakeholder values. Tatar emphasizes that such conflicting values cannot often be fully resolved, but "only handled via compromise" [81]. Prior HCI research has used the framework, or its variants, to study or synthesize design requirements in a range of domains, including health promotion applications [49], learning analytics dashboards [80], online behavioral therapy experiences [41], and wearable technology to support children with ADHD [19]. As designing app-based assessments for scaling behavioral health screening outside clinical settings inherently necessitates balancing several benefits and drawbacks, we found the design tensions framework suited to our analysis. Section 6.1 highlights three major design tensions emerging from our analysis. Finally, we integrated our findings to offer recommendations for balancing the identified design tensions and designing usable and scalable screening technologies in Section 6.2.

3.4 Positionality Statement

This work stems from the authors' long-standing research interests in technology and mental health. The authors have previously developed tools for a range of mental health and well-being applications, including for children's behavioral health. Some authors have experienced mental health issues and/or have others they care for who have experienced mental health hardships. Three of the authors have experience working with preschool-aged children and their families in clinical settings as registered behavior technicians or clinical psychology trainees, while one author has experience working in special education classrooms with children with autism spectrum disorder. Three of the authors are parents of children who are preschool-aged or older. Half the authors identify as women and the other half as men. The authors' self-described racial identities include, in alphabetical order, Asian (Bangladeshi origin), Asian/Indian, Black Latina, Latinx of Indigenous (American) and Chinese descent, and white. All authors reside in the United States.

4 FINDINGS: EXISTING CONCERNS AND HOW SCREENING TOOLS CAN ADDRESS THEM

In this section, we describe our findings highlighting the lived experience of child mental health practitioners and parents of preschool-aged children in terms of their concerns about children's behavioral health, gaps in current screening practices, and perceived benefits and challenges in deploying and using at-home mental health screening apps to support their clinical or parenting practices. As described in Section 3.3, we fuse insights from both stakeholder groups to derive broad themes highlighting their collective perspectives. Quotes from clinicians are identified by the pseudonyms C1 through C25 and those from parents by P1 through P25.

4.1 Stakeholder Concerns and Gaps in Pediatric Behavioral Screening

4.1.1 Parents lack scaffolding for identifying concerns and seeking help. A common theme that emerged from discussions with parents, both those with and without specific concerns about their child's behavior, was not knowing what to expect in terms of behavioral development at a certain age and when to seek help. P4b described the need for a developmental equivalent to growth charts, saying,

> *"As a parent, you will be given those developmental charts. So developmentally, like by three months, what he is expected to do. By one year, he will be able to stand with support ... For psychological or mental things, we don't have that chart. For frustration, social behavior, or tantrums – those are not well documented. There is nothing about what to expect at what age for psychological development."* (P4b)

As a result, parents' help-seeking behavior is largely driven by their intuitions (*"if we feel something is off, then we talk to the doctors"* (P4b)) as well as their willingness and comfort in approaching others for help (*"it's like, I don't know, should I ask for help maybe?"* (P12)) instead of the more standardized approaches utilized in physical medicine. P13 also emphasized that the lack of concrete information prevented parents from successfully advocating for their child:

"[In many public programs] there's some resources available but generally not enough to go all the way around. And so there's like a gatekeeping process where they basically try to determine who should really get these resources, and navigating that process can be really, really difficult." (P13)

In terms of overcoming these challenges, most parents reported reaching out to family and friends, doctors, teachers, or community support programs and organizations to discuss their concerns and obtain more information about their child's development, in addition to accessing information from books or the internet. However, they expressed frustration at being subjected to an information overload (*"the amount of information out there is overwhelming and it's hard to parse through it to find what's actually useful"* (P10)) or not being able to receive confirmation of typical development even if they do not have existing concerns (*"for the most part I think it's just the typical kid frustration for them. [But] it'd be good to know, it's always good."* (P16)).

4.1.2 Limited resources available for tracking and supporting emotional development. In addition to the lack of a social-emotional development scaffold to help identify concerns, stakeholders emphasized the paucity of resources for supporting families in navigating existing concerns. P15 described struggling to find the right strategies to support her child who was prone to emotional outbursts and a short temper:

"I have a pretty good resource connection because I do have an education background. So I developmentally know what should be happening and who I should talk to – doctor-wise or other teachers and behavior specialist-wise, but nothing's quite clicked yet." (P15)

While some parents struggled to find actionable ideas, others would seek reassurance from healthcare providers as they navigated behavioral concerns (*"what I want to hear is like, "hey this is fine ... it can be fixed by doing this"."*(P12)). However, practitioners themselves struggled to accurately diagnose or track children's symptoms over time, with C12 stating,

"I often work in settings that have a dearth of resources available for both families and clinicians ... While there are many free narrowband measures available, there are few reliable broadband measures that give the same level of detail as a BASC or CBCL." (C12).
Authors' note: BASC (Behavior Assessment System for Children; [70]) and CBCL (Child Behavior Checklist; [1]) are both commonly used proprietary broadband assessment scales for children.

Similarly, clinicians also expressed the need for *"measures that can be administered at higher frequency (e.g., daily)"* (C2) in order to continually track children's behavior.

4.1.3 Behavioral screening is based on incomplete data. Another common challenge encountered by mental health practitioners working with young children was the limited availability of data required for diagnosis and treatment. Clinicians reported wanting access to clinical interview data from both parents instead of just one (C9) and from biological parents in case of adoption or foster care scenarios (C19). They also expressed a need for school and community data, including classroom observations, teacher interviews, school records, and information about how the child interacts with other adults (C1, C7, C11, C13, C14, C15, C19, C24). Further, the diagnostic process often lacks access to behavioral data (C8, C22) and real-world observations outside the clinic (C10, C17).

Additionally, behavioral diagnoses rely on brief observations of the child in a clinical setting. P1 explains how this misses important context and thus decreases the reliability of diagnosis:

"Anyone that is trying to know the kid should not spend a couple of hours in a meeting to understand him. So a couple of days, sometimes a couple of weeks, might get you like, a day-to-day or an overall idea of how these kids behave. But in a meeting, with kids that you do not know – for instance a psychologist that spent like an hour with a kid – the kid will be shy, he does not know her/him. And he will not express his feelings as a grown-up would. It will take weeks." (P1)

These limitations mean that even families who are able to access mental health services often fail to receive the highest level of care due to missing data or limited contextual information.

4.2 Potential Benefits of Screening Tools

4.2.1 Increasing access to mental health services. Both clinicians and parents felt that app-based, at-home screening tools had the potential to *"increase access to care for underserved communities"* (C2) and reach numerous families as long as they *"have a phone/tablet and WiFi"* (C7). Stakeholders appreciated the convenience of use (C2) and noted that app-based screening could be particularly *"effective for low socioeconomic status families"* (C4) and that it was *"easy and convenient enough, and doesn't seem super invasive"* (P17).

Additionally, parents valued screening tools as a first step towards seeking professional mental healthcare. P13 described wanting to test their child at home by saying,

"I would say if we gave him one of these [tools] and it was like, "Oh we think your kid, [there's] a high percentage chance he has this", yeah, we'll probably follow up with a doctor, like, pronto. Yes, that would be enough to get action for sure. ... We'll probably give it a try because we'd be really interested to see." (P13)

4.2.2 Augmenting existing data and clinical practices. Clinicians appreciated that app-based screening tools were able to *"provide a different type of information to that [clinicians are] able to collect"* (C17), and specifically highlighted *"the passive ability to gather information"* (C15) as a key benefit. Parents also felt that such tools could supplement regular healthcare appointments by providing additional insight into the child's development:

"I just think [screening tools] are something good because maybe with [these tools] you can get some preliminary information. Which maybe a doctor can't identify when you are at a checkup. So it can be something which is in addition to a routine checkup of a child." (P5)

Parents similarly noted that *"having a huge database of information to compare one person to, could potentially result in a more reliable prediction or diagnosis than just one health care provider"* (P6) and

were interested in using screening tools *"in conjunction with, like, an actual provider"* (P6).

Clinicians also listed additional data streams that could be collected, or functions that could be supported, via app-based screening tools. C7 wanted to see future tools that could provide *"the ability [for] the mental health professional to observe the child while engaged in the task and/or the ability to provide live support to the parent/caregiver during/after the task"* (C7). Others wanted to use such tools to measure parent-child interaction (C1, C2), caregivers' frustration tolerance (C12), or longitudinal progress (C6, C7).

4.2.3 Making screening more frictionless for children. Another important benefit of app-based screening tools, stakeholders noted, was that they could make mental health screening more engaging, natural, and frictionless for young children. Participants thought game-based screening would be *"creative and engaging for young children"* (C6), especially since tablet-based games with colors, matching objects, pretend play, and cartoon characters would be relevant to children's existing interests (*"I think [child would] love it. It's adorable and … he watches Daniel Tiger [children's show]. And he loves the Baker Aker [character]. So it would totally be in his wheelhouse"* (P10)).

Parents and clinicians also felt that game-based or other naturalistic modes of screening would be particularly beneficial for children who would feel uncomfortable with traditional assessments. C12 explained,

> *"It is sometimes difficult to observe disruptive behaviors when children come to a clinic and meet individuals with whom they are unfamiliar. This application would potentially remove demand characteristics at play within a clinic and capture the child behaving in their natural environment."* (C12)

Parents agreed that such tools would be very helpful to children, *"more than the formal appointments that make kids sometimes scared"* (P11). P13 also said his child would *"not sit there and have a long detailed conversation with you about how he feels about these things. He's quickly going to get bored"* (P13), and engaging, naturalistic assessments would be a useful alternative to test such children.

4.2.4 Supporting the parenting experience. Parents articulated how screening tools would support them through their own struggles by taking some pressure off of the parenting experience. P15 explained,

> *"As a baby, for [child], I did [an app with] different things you can do with your baby for the first year … And I found that particularly helpful for me because I had my own mental health stuff going on at that time with postpartum. So not having to think about it but felt like, "All right, still getting stuff done, I'm still doing a good job. I'm still getting answers". I think that's where [screening apps] would fall under".* (P15)

In addition to letting parents automatically engage in the screening process without explicit effort, screening apps could also help address the "when-to-worry" problem (*"if there was something related to behavior, which might be unusual, which we observe … And if that is something which is widely observed – we would probably get to know it from such [apps].* (P5)). Screening tools can also provide

parents additional insight in terms of *"something to go to the doctor with … useful information for the follow-up"* (P13).

4.3 Challenges in Deploying Screening Tools

4.3.1 Integrating with traditional mental health services and enabling opportunities for follow-up. While stakeholders valued at-home screening tools as complementary to existing mental health services, some participants also expressed concerns that such tools might *"take away from scientifically validated in-person assessments in sterile clinic environment"* (C1) or *"override clinical training/judgment"* (C9). Over-reliance on app-based tools could also disproportionately affect populations who are less likely to be able to access traditional services. P6 commented,

> *"I'm worried that technology which is cheap and broadly available is disproportionately used by lower-income individuals and individuals that have had difficulty accessing traditional services, in lieu of improving accessibility to those services."* (P6)

C10 similarly felt that *"a low-income or otherwise busy family who cannot make a diagnostic interview is also unlikely to follow-up for care"* (C10) and said they would want to *"see data that suggests the app actually leads to clinical follow-up in high-risk populations"* (C10). Others suggested that the app itself could include a *"treatment related intervention"* (C19) or actionable next steps for parents on *"what to do about it, if there's problems that come up"* (P18).

4.3.2 Building stakeholder trust in screening technologies. Another challenge in deploying screening tools for young children lies in establishing stakeholder trust in such technologies. Some participants shared their distrust of mental health technology due to negative perceptions of existing commercial products (*"I've heard a lot of horror stories about [redacted online therapy service provider]. And like, it's honestly kind of scared me off [of] this sort of tech in general"* (P6)). Parents were also concerned about the implications of apps inferring sensitive characteristics of their child's health, and how it might impact their future. P6 added,

> *"I'm worried that if there's something forming psychological profiles of children, that's going to follow them their entire lives. I could imagine that it will leak into the education systems, like, somebody who tests poorly on this device is going to get red-shirted."* (P6)

Stakeholders also had questions about child and parental consent (*"Will children always seek parental consent before using the app?"* (C10)) and concerns about the lack of situational context (C4, C15) as well as the accuracy of screening tools (*"machine learning can be imperfect, and can misidentify or miss at-risk children"* (C10)). Moreover, parents stated that they would need explanations for the tool's outcomes to be able to take it seriously:

> *"For any app to gain trust, like, if I give him a game and the only thing the app says is "something is wrong with your child", I wouldn't accept that. Make me understand why you are right."* (P4b)

Other parents echoed this sentiment, saying feedback in layperson terms (*"something like, you know, "your child spent a particularly long time doing this, or reacted this way" … whatever you can do*

to communicate what it is you think you found" (P13)) would help build trust in the tool.

4.3.3 Minimizing potential for harm. Lastly, participants emphasized the need to account for safety, privacy, and ethical considerations when deploying screening technologies. In terms of child safety, parents were worried that tools could be *"too involved"* (P13) or misleading for the child (*"it might be hard for me to explain [outcomes] if there is no lesson in it"* (P3)). Other parents were wary of introducing or increasing their child's screen time or device usage (multiple parents), or enabling unrestricted access to the Internet (P15), by using app-based tools. Stakeholders also felt that children might eventually get bored or frustrated with repeated use of screening tools (*"I think that you'd have to change [the game] up a lot because I think they would get bored"* (P14)), or that they may cause undue stress for children and/or parents. P1 elucidated the latter point, saying,

> "I think in the US, there is too much of a focus on diagnosing kids, categorizing them ... I don't think that is necessary ... And even if the kids start showing any slight symptoms, of any issue, they try to always put way too much pressure on the parents and on [the child]. That might not be, uh, motivating for people to go seek help". (P1)

Other parents seconded this, while P3 added that *"if [the tool] gives like a lot of false positives, then it just becomes more stressful"* (P3). In addition to ensuring that tools are accurate, engaging, and safe for children to use, developers also need to focus on data storage and privacy protections. Stakeholders emphasized the need for *"really rigorous data protections, including not harvesting and storing anything that isn't totally vital"* (P6) and guaranteeing *"confidentiality"* (C2), and *"privacy"* (C4, C13, P7).

5 FINDINGS: DIFFERENCES IN PERSPECTIVES ACROSS STAKEHOLDER GROUPS

While Section 4 highlights the broad collective perspectives of various stakeholders toward app-based screening tools and illustrates how clinicians and parents largely agree on many facets, it is important to also understand where the two groups may have opposing views. Below, we describe some key differences emerging from our analysis of clinician and parent responses.

5.1 Intentionally Inducing Frustration

A characteristic feature of the EarlyScreen app [44], which has been used as a design probe in this work, is that it induces frustration in children in order to record how they respond to negative emotional stimuli. While this approach may not be a necessary component of all app-based screening tools, it is worth discussing stakeholders' views toward intentionally inducing frustration. Clinicians in particular were wary of this practice, noting concerns about *"managing a child's reaction to the frustration scenario remotely or leaving the parent/caregiver to deal with whatever happens without live support"* (C7). While clinicians acknowledged that *"this [frustration] often happens naturally and this certainly occurs when implementing practices like exposure with response prevention"* (C12) in routine behavioral healthcare settings, they were nevertheless concerned

that *"parents may be very wary of using [EarlyScreen] as it induces frustration without clinical supervision"* (C8).

Interestingly, parents did not share this concern – while they noted that repeated frustration could result in children not completing or repeating the task frequently (P15, P17, P24), parents were highly interested in the possibility of understanding how their child would react to deliberate and controlled frustration. For instance, P4b felt that this approach *"has a metric"* as opposed to *"right now, where if he wants something and we don't give him that, then comes the frustration or anger or tantrum"* (P4b). P15, who mentioned concerns about her child's irritability, said,

> "I like the idea of it purposely being frustrating to see how they would react, and feel like [child] would be a great candidate for that". (P15)

Parents further viewed the frustration-inducing game as an opportunity to discuss emotions with children and help them develop emotion regulation capabilities. P22 mentioned that such tools can *"help [child] learn or get the feelings out"* (P22), while P9 said,

> "Children tend to be more frustrated and angry when things don't work out in their support ... I would like to create a feeling of patience in them. So if this app helps reach that, and helps in assessing my child, that will be great". (P9)

5.2 Apps and Screen-based Interaction Format

Parents were overwhelmingly more apprehensive about the screen-based assessment medium in EarlyScreen, with over half of them raising concerns about screen time, device usage, or potentially unrestricted access to the Internet while completing the screening. P8 said she probably wouldn't use EarlyScreen, adding:

> "Me and a lot of the parents I know are super against screen time as much as possible. So we really limit it, even if it's like a game or like [educational content] and stuff, I prefer books and paper media". (P8)

P3 shared these concerns about screen time and further added,

> "I'm okay, or more okay, with TV and [gaming] consoles. I'm less comfortable with kids playing with phones and iPads because those devices access the internet very easily. So they are more exposed to potentially harmful information. Also, there are a lot of micro-transactions in iPad games and iPhone games. It's very easy for kids to purchase something by accident. And also I think a lot of the iPad games and the phone games are just designed to be very addictive compared to TV shows". (P3)

Parents also felt that the novelty of screen-based tasks might encourage children to ask to do them more often than is necessary, especially if their screen time was usually very limited (P2, P4, P9, P11, P21). Multiple parents (P3, P8, P9) said they would be more amenable to home-based screening if it did not involve a screen, while one (P7) wanted more data on how using the app might impact children in the long term.

In contrast, none of the clinicians surveyed expressed any concerns about screen/device exposure. Clinicians were more excited about the gamified format of the assessment, the potential ease of

installation and use, and the increased access to screening afforded by this format (see Section 4.2.1).

5.3 Integration Within Healthcare Settings

Another important distinction between clinicians and parents was their perspective on the integration of app-based screening tools into the traditional mental health care ecosystem. While both groups cautioned against an over-reliance on such tools at the cost of clinical or parental judgment or improving access to health services, parents still saw value in at-home screening tools that were situated outside healthcare settings. P16 explained how screening tools were valuable for making sure there was indeed nothing to be concerned about (*"Better to know more information than not, I think."* (P16)).

Parents also felt that at-home screening tools would enable them to track children's socio-emotional development relative to peers, help identify concerns early and ascertain the level of risk associated with them, and empower them in communicating with health care providers. P20 hoped they would provide *"a little more info before you go to see someone outside of the home"* (P20).

Relatedly, while clinicians were concerned about the potential lack of follow-up, most parents said that they would immediately seek help from a medical professional if an at-home assessment identified a possible disorder or concern:

> *"I would say if we gave him one of these things and it was like, "Oh we think your kid – a high percentage chance he has this", we'll probably follow up with a doctor like pronto. Yes, that would be enough to get action for sure".* (P13)

6 DISCUSSION

6.1 Design Tensions in Early Childhood Mental Health Screening Tools

As evidenced by our findings in Sections 4 and 5, designing app-based mental health screening tools for children inherently calls for balancing several key considerations from different stakeholders. We examine these factors through the lens of Tatar's design tensions framework [81], identifying three conflicting pairs of requirements that are elucidated upon here. Note that these tensions should not be viewed as pitting stakeholder groups against each other (parents vs. clinicians) but instead as an expansive set of *design* considerations emerging from both agreements and disagreements in a value-sensitive multi-stakeholder analysis. We make design recommendations for balancing these design tensions in Section 6.2. The design tensions and recommendations are summarized in Table 2.

6.1.1 Need for rich observational data vs. limiting burden and exposure. Both mental health practitioners and parents expressed a desire for screening tools to capture behavioral data in a variety of contexts and often longitudinally. This has three clear benefits – first, families can obtain more reliable and trustworthy predictions through repeat testing, minimizing the impact of temporary contextual factors on assessment outcomes. Second, it allows both parents and clinicians to monitor behavior progression over time and in response to interventions. Finally, clinicians can obtain rich, naturalistic data that would not be available in an outpatient clinical setting. In addition to longitudinal data, parents expressed an

interest in observing children's response to controlled frustration in EarlyScreen, motivating future tools that can deliver regulated measures of stimuli and record the resulting behavior.

However, the repeated use of screening tools comes at a cost. Parents may want to avoid exposing their child to tools repeatedly (e.g., in order to minimize screen time) and families may find recurring assessments time-consuming or burdensome. Longitudinal use also requires considering practice effects and keeping the child engaged over time. Inducing negative affect to observe responses necessitates deliberation on the quantity and quality of stimulus to avoid adverse effects for the child and to ensure that families can support any outcomes in the moment without clinical intervention. Screening tools therefore need to balance the reliability and utility of data against usability concerns.

6.1.2 Frictionless screening vs. over-reliance on tools. App-based and gamified screening tools can be more child-friendly than traditional methods of assessment, especially for children who fear doctors, get easily distracted, or are reserved when interacting with strangers. Similarly, they can increase parental convenience and empower parents by providing useful information about their child's wellbeing and their development relative to peers [88]. Lastly, at-home screening tools can provide parents with a first-step diagnosis, spur them toward seeking clinical follow-up, and help communicate their concerns to a clinician.

Nevertheless, both clinicians and parents expressed concerns about overly relying on the outcomes of screening tools. It can be detrimental to the child's health if families are unable to follow up on screening outcomes and get the required help. Screening technologies should also not result in loss of parental intuition or reliance on their experiential knowledge, and providers should not let them override their clinical judgment [10]. Tools should instead provide stakeholders with clear pathways to support a complementary caregiving process [45].

6.1.3 Scaling beyond the clinic vs. minimizing contamination and harm. Screening tools can dramatically reduce barriers to obtaining a behavioral diagnosis for many families, eliminating waiting times and making screening more affordable and accessible [8]. However, they may potentially take away from the sterile examination setting of a clinic, making screening subject to a wide range of factors (e.g., child's mood and surroundings, parental influence, presence of siblings and pets, device variability, etc.) that are difficult to standardize across children and families in home settings [6, 91]. Furthermore, screening tools would need to account for privacy and data storage considerations [33] – traditional healthcare services often have existing mechanisms in place to address these concerns.

6.2 Design Implications and Recommendations

Based on our findings and the design tensions identified in the previous subsection, we now discuss design opportunities for building scalable mental health screening technologies for young children.

6.2.1 Explore Novel Assessment Mediums. A major concern for parents in potentially using app-based screening tools was the addition of screen time into their children's lives. We call upon designers to address this by (i) exploring app-based assessment modalities that avoid media exposure or utilize passive sensing (e.g.,

Table 2: Design Tensions and Recommendations for Building Mental Health Screening Tools for Children

Design Tension	Considerations	Recommendations
Need for rich observational data vs. limiting burden and exposure (Section 6.1.1)	Clinicians and parents want tools to observe a child's behavior longitudinally and in response to fixed stimuli, both to collect deeper insights and to build trust. However, administering such tools may cause additional burdens for families.	Explore screen-free assessment mediums for longitudinal monitoring and allow parents to customize stimuli for their child (Sections 6.2.1 and 6.2.2).
Frictionless screening vs. over-reliance on tools (Section 6.1.2)	While gamified, at-home screening tools can be beneficial for children who do not want to visit doctors or engage with traditional screening methods, solely relying on such tools can be detrimental.	Build engaging screening experiences tailored to children's interests while incorporating clear next steps for parents to follow (Sections 6.2.3 and 6.2.4).
Scaling beyond the clinic vs. minimizing contamination and harm (Section 6.1.3)	Digital tools have the potential to highly scale mental health screening, but observation and inference outside sterile clinical settings may lead to data contamination and raise privacy concerns.	Integrate measures of contextual variability into screening algorithms and add protections for sensitive data and outcomes (Sections 6.2.5 and 6.2.6).

audio prompts and recording, wearable devices [88], etc.) and (ii) incorporating more naturalistic methods beyond app-based sensing. Developers can draw inspiration from prior work that has leveraged assessment mediums such as instrumented building blocks [89] and interactive non-anthropomorphic robots [15].

6.2.2 Enable Parental Control Over Stimuli. The EarlyScreen app uses negative stimuli to deliberately induce frustration, which clinicians worried could trigger undesirable consequences for children's mental state. Screening tools can instead attempt to observe naturally occurring instances of frustration. However, this would require observation over much longer timescales if such behaviors are infrequent, leading to increased burden and screen time. As an alternative, we envision future tools observing responses to discrete levels of stimulus exposure where parents can determine the acceptable grade of stimulation. For example, in a variation of EarlyScreen's frustration-inducing game, parents could adjust the intensity of negative feedback that they feel would be tolerable for their child as well as the frequency of positive and negative feedback before beginning the assessment.

6.2.3 Build for Engagement. For repeated assessment using screening tools to be successful, it is necessary to keep children engaged and motivated to use the tool without getting bored or overly frustrated. Prior work has shown that children engage with mobile apps through sensory experiences such as touching, looking, and listening, as well as emotional and verbal expressions in response to the app [66]. Screening tools in any chosen medium should aim to replicate such engagement by designing experiences that incorporate children's real-life interests. Such tools should also cater to children with a broad range of interests and abilities (e.g., very young children or children with behavioral issues may be less likely to engage with overly complicated narratives). Rigorous user testing or participatory design approaches can help develop either broadly appealing or customized experiences for different kids [38].

6.2.4 Enable Pathways for Future Care. Prior work has discussed the importance of child tracking technologies to use the data they collect to empower parents and support their caregiving practices, instead of promoting an over-reliance on predicted outcomes at the cost of parental or clinical judgment [45]. Mental health screening tools, in particular, should delineate clear next steps, including (i) informing parents of potential risks and normative comparisons across peers [62], (ii) providing consolidated information to help them articulate concerns when seeking help [30], (iii) connecting them to sources of support within and outside the traditional behavioral care system, (iv) facilitating relationship-building and information sharing across care teams [71], and (v) providing reassurance at a time of distress and supporting them in navigating their own mental health [4].

6.2.5 Controlling for Context and Measuring Variability. Screening tools that collect assessment data outside controlled clinical settings also need to account for contextual variability and potential contamination due to environmental factors. One approach to limit variability is to observe the child during a particular, focused activity (such as the frustration-inducing game in EarlyScreen). Screening algorithms can also explicitly record and utilize variability such as behavioral context [7] and parent-child interactions [36] in order to make predictions. This would also serve to provide additional metrics for clinicians who subsequently review the data to assess the child.

6.2.6 Ensure Strong Data Protections. It is important for screening tools to incorporate strict data privacy measures in order to protect sensitive outcomes pertaining to vulnerable populations and to build stakeholder trust. Some measures to this end include only recording absolutely essential information, not transmitting data or predicted outcomes off of the user's device, and expunging records that are no longer needed. Prior work has also demonstrated that caregivers may be more comfortable sharing data from in-home child monitoring technologies if they were anonymized at the point of collection and if parents were able to access, review, and possibly censor recordings before sharing [27].

6.3 Contributions to HCI and Child Mental Health Research

As described in Section 2, there has been a growing interest in the HCI and digital health communities to augment traditional mental health services using behavioral screening tools to identify problem behaviors. However, building useful and scalable screening tools, especially for young children, necessitates understanding the unmet needs in the current mental health care landscape from the perspective of various participants in the caregiving process as well as determining to what extent these needs can be fulfilled by at-home screening technologies.

Our work is the first to investigate these questions, focusing specifically on app-based, at-home screening tools for behavioral screening in preschool-aged children. It makes the following contributions to the broader HCI community interested in developing screening tools for young children: (i) we present empirical findings on the attitudes of mental health practitioners and parents concerning at-home screening tools, (ii) we highlight nuances in stakeholder views by disentangling the agreements, conflicting priorities, and overarching design tensions emerging from our multi-stakeholder inquiry, and (iii) we offer concrete recommendations for designing future child mental health screening tools, distilling insights from both our own analyses as well as prior research.

6.4 Limitations and Future Work

Our work examined the perspectives of two key stakeholders in the child mental health care system – clinicians and parents – to inform the design of scalable screening technologies for children. Our studies focused on child and adolescent mental health practitioners across the United States as well as a predominantly White/European-American/Asian sample of parents of preschool-aged children. As such, further research is required to investigate whether our findings generalize to other geographical and demographic contexts. The perspectives of children themselves as users of these screening tools were also not examined in this work due to the nature of our inquiry. We encourage future studies to involve children as stakeholders in the design of screening technologies through age-appropriate participatory design approaches [23, 38, 40]. Lastly, our decision to use EarlyScreen as a speculative probe may have also biased participant responses towards a particular class of screening tools (app-based, data-driven assessments with active participation). However, we felt it was necessary to utilize an example screening tool as a provocation in order to both elicit feedback of a speculative nature [2, 16] and to familiarize participants who may not have used or even come across such tools before. This was confirmed during the parent interviews, where most parents reported being unfamiliar with at-home behavioral screening technology. Future research could elicit stakeholder feedback on other kinds of screening tools, and triangulate them against the findings presented here.

7 CONCLUSION

This work sought to understand the perspectives of mental health practitioners and parents of preschool-aged children towards digital mental health screening technologies. We reported on a survey of clinicians and interviews with parents in the United States, focusing on uncovering the barriers they currently face in identifying and seeking help for behavioral issues in young children and how screening tools could support their caregiving practices. Our work uncovered three key tensions relevant to the design of screening tools: (i) the need for rich observational data vs. limiting burden and exposure, (ii) frictionless screening vs. over-reliance on tools, and (iii) scaling beyond the clinic vs. minimizing contamination and harm. Finally, we offered recommendations for developers of future tools based on insights from multiple stakeholders and delineated opportunities for future research.

ACKNOWLEDGMENTS

REFERENCES

[1] Thomas M Achenbach and Leslie A Rescorla. 2000. *Manual for the ASEBA preschool forms and profiles*. Vol. 30. Burlington, VT: University of Vermont, Research center for children, youth, & families.

[2] Daniel A Adler, Emily Tseng, Khatiya C Moon, John Q Young, John M Kane, Emanuel Moss, David C Mohr, and Tanzeem Choudhury. 2022. Burnout and the Quantified Workplace: Tensions around Personal Sensing Interventions for Stress in Resident Physicians. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (Nov. 2022), 1–48. https://doi.org/10.1145/3555531

[3] Dhruv Agarwal, Srishti Agarwal, Vidur Singh, Rohita Kochupillai, Rosemary Pierce-Messick, Srinivasan Iyengar, and Mohit Jain. 2021. Understanding Driver-Passenger Interactions in Vehicular Crowdsensing. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 1–24. https://doi.org/10.1145/3479869

[4] Ebtisam Alabdulqader, Katy Stockwell, Kyle Montague, Dan Jackson, Andrew Monk, Lindsay Pennington, Roisin McNaney, Stephen Lindsay, Ling Wu, and Patrick Olivier. 2022. Understanding the Therapeutic Coaching Needs of Mothers of Children with Cerebral Palsy. In *Designing Interactive Systems Conference (DIS '22)*. ACM, 1788–1801. https://doi.org/10.1145/3532106.3533576

[5] Jordi Alonso, Zhaorui Liu, Sara Evans-Lacko, Ekaterina Sadikova, Nancy Sampson, Somnath Chatterji, Jibril Abdulmalik, Sergio Aguilar-Gaxiola, Ali Al-Hamzawi, Laura H. Andrade, Ronny Bruffaerts, Graça Cardoso, Alfredo Cia, Silvia Florescu, Giovanni de Girolamo, Oye Gureje, Josep M. Haro, Yanling He, Peter de Jonge, Elie G. Karam, Norito Kawakami, Viviane Kovess-Masfety, Sing Lee, Daphna Levinson, Maria Elena Medina-Mora, Fernando Navarro-Mateu, Beth-Ellen Pennell, Marina Piazza, José Posada-Villa, Margreet ten Have, Zahari Zarkov, Ronald C. Kessler, and Graham Thornicroft. 2018. Treatment gap for anxiety disorders is global: Results of the World Mental Health Surveys in 21 countries. *Depression and Anxiety* 35, 3 (Jan. 2018), 195–208. https://doi.org/10.1002/da.22711

[6] Alaa Alslaity, Gerry Chan, Rita Orji, and Richard Wilson. 2022. Insights From Longitudinal Evaluation of Moodie Mental Health App. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '22)*. ACM, 1–8. https://doi.org/10.1145/3491101.3519851

[7] Riku Arakawa, Karan Ahuja, Kristie Mak, Gwendolyn Thompson, Sam Shaaban, Oliver Lindhiem, and Mayank Goel. 2023. LemurDx: Using Unconstrained Passive Sensing for an Objective Measurement of Hyperactivity in Children with no Parent Input. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 2 (June 2023), 1–23. https://doi.org/10.1145/3596244

[8] Patricia A. Areàn, Kien Hoa Ly, and Gerhard Andersson. 2022. Mobile technology for mental health assessment. *Dialogues in Clinical Neuroscience* 18, 2 (June 2022), 163–169. https://doi.org/10.31887/dcns.2016.18.2/parean

[9] American Psychological Association. 2022. *Society of Clinical Child and Adolescent Psychology*. Retrieved December 28, 2022 from https://sccap53.org/

[10] Helen L Ball and Alice-Amber Keegan. 2022. Digital health tools to support parents with parent-infant sleep and mental well-being. *npj Digital Medicine* 5, 1 (Dec. 2022), 1–4. https://doi.org/10.1038/s41746-022-00732-4

[11] Jennifer Beecham. 2014. Annual Research Review: Child and adolescent mental health interventions: a review of progress in economic studies across different disorders. *Journal of Child Psychology and Psychiatry* 55, 6 (Feb. 2014), 714–732. https://doi.org/10.1111/jcpp.12216

[12] Ayelet Ben-Sasson, Eli Ben-Sasson, Kayla Jacobs, Elisheva Rotman Argaman, and Eden Saig. 2019. Evaluating Expert Curation in a Baby Milestone Tracking App. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, 1–12. https://doi.org/10.1145/3290605.3300783

[13] Katherine Bessière, A. Fleming Seay, and Sara Kiesler. 2007. The Ideal Elf: Identity Exploration in World of Warcraft. *CyberPsychology & Behavior* 10, 4 (Aug. 2007), 530–535. https://doi.org/10.1089/cpb.2007.9994

[14] Rhys Bevan Jones, Paul Stallard, Sharifah Shameem Agha, Simon Rice, Aliza Werner-Seidler, Karolina Stasiak, Jason Kahn, Sharon A. Simpson, Mario Alvarez-Jimenez, Frances Rice, Rhiannon Evans, and Sally Merry. 2020. Practitioner review: Co-design of digital mental health technologies with children and young people. *Journal of Child Psychology and Psychiatry* 61, 8 (June 2020), 928–940. https://doi.org/10.1111/jcpp.13258

[15] Laura Boccanfuso, Erin Barney, Claire Foster, Yeojin Amy Ahn, Katarzyna Chawarska, Brian Scassellati, and Frederick Shic. 2016. Emotional robot to examine different play patterns and affective responses of children with and without ASD. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 19–26. https://doi.org/10.1109/hri.2016.7451729

[16] Kirsten Boehner, Janet Vertesi, Phoebe Sengers, and Paul Dourish. 2007. How HCI interprets the probes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI07)*. ACM, 1077–1086. https://doi.org/10.1145/1240624.1240789

[17] Elizabeth H Bringewatt and Elizabeth T Gershoff. 2010. Falling through the cracks: Gaps and barriers in the mental health system for America's disadvantaged children. *Children and Youth Services Review* 32, 10 (Oct. 2010), 1291–1299. https://doi.org/10.1016/j.childyouth.2010.04.021

[18] Alice S. Carter, Margaret J. Briggs-Gowan, and Naomi Ornstein Davis. 2004. Assessment of young children's social-emotional development and psychopathology: recent advances and recommendations for practice. *Journal of Child Psychology and Psychiatry* 45, 1 (Jan. 2004), 109–134. https://doi.org/10.1046/j.0021-9630.2003.00316.x

[19] Franceli L Cibrian, Kimberley D Lakes, Arya Tavakoulnia, Kayla Guzman, Sabrina Schuck, and Gillian R Hayes. 2020. Supporting Self-Regulation of Children with ADHD Using Wearables: Tensions and Design Challenges. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. ACM, 1–13. https://doi.org/10.1145/3313831.3376837

[20] Victoria Clarke and Virginia Braun. 2013. *Successful qualitative research: A practical guide for beginners*. Sage publications ltd. 1–400 pages.

[21] P Cohen and C S Hesselbart. 1993. Demographic factors in the use of children's mental health services. *American Journal of Public Health* 83, 1 (Jan. 1993), 49–52. https://doi.org/10.2105/ajph.83.1.49

[22] Nicholas Deveau, Peter Washington, Emilie Leblanc, Arman Husic, Kaitlyn Dunlap, Yordan Penev, Aaron Kline, Onur Cezmi Mutlu, and Dennis P Wall. 2022. Machine learning models using mobile game play accurately classify children with autism. *Intelligence-Based Medicine* 6 (2022), 100057. https://doi.org/10.1016/j.ibmed.2022.100057

[23] Allison Druin. 2002. The role of children in the design of new technology. *Behaviour & Information Technology* 21, 1 (Jan. 2002), 1–25. https://doi.org/10.1080/01449290110108659

[24] Nancy Eisenberg and Richard A Fabes. 1992. Emotion, regulation, and the development of social competence. (1992).

[25] Daniel Fatori, Giovanni Salum, Alexander Itria, Pedro Pan, Pedro Alvarenga, Luis Augusto Rohde, Rodrigo Bressan, Ary Gadelha, Jair de Jesus Mari, Maria Conceição do Rosário, Gisele Manfro, Guilherme Polanczyk, Euripedes Constantino Miguel, and Ana Soledade Graeff-Martins. 2018. The economic impact of subthreshold and clinical childhood mental disorders. *Journal of Mental Health* 27, 6 (April 2018), 588–594. https://doi.org/10.1080/09638237.2018.1466040

[26] Jessica L Feuston and Jed R Brubaker. 2021. Putting Tools in Their Place: The Role of Time and Perspective in Human-AI Collaboration for Qualitative Analysis. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 1–25. https://doi.org/10.1145/3479856

[27] Laurel A. Fish and Emily J. H. Jones. 2021. A survey on the attitudes of parents with young children on in-home monitoring technologies and study designs for infant research. *PLOS ONE* 16, 2 (Feb. 2021), e0245793. https://doi.org/10.1371/journal.pone.0245793

[28] Theresa M. Fleming, Derek de Beurs, Yasser Khazaal, Andrea Gaggioli, Giuseppe Riva, Cristina Botella, Rosa M. Baños, Filippo Aschieri, Lynda M. Bavin, Annet Kleiboer, Sally Merry, Ho Ming Lau, and Heleen Riper. 2016. Maximizing the Impact of e-Therapy and Serious Gaming: Time for a Paradigm Shift. *Frontiers in Psychiatry* 7 (April 2016), 65. https://doi.org/10.3389/fpsyt.2016.00065

[29] Radhika Garg. 2021. "Its Changes so Often": Parental Non-/Use of Mobile Devices While Caring for Infants and Toddlers at Home. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 1–26. https://doi.org/10.1145/3479513

[30] Doğa Gatos and Asim Evren Yantaç. 2020. "Oxygen Mask": Understanding How Autism Parents Seek Support. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20)*. ACM, 1–11. https://doi.org/10.1145/3419249.3420153

[31] Lorna Gibson and Vicki L Hanson. 2013. Digital motherhood: how does technology help new mothers?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, 313–322. https://doi.org/10.1145/2470654.2470700

[32] Hannah Gilbert, Ling Qin, Dandan Li, Xuehua Zhang, and Stuart J Johnstone. 2016. Aiding the diagnosis of AD/HD in childhood: Using actigraphy and a continuous performance test to objectively quantify symptoms. *Research in Developmental Disabilities* 59 (Dec. 2016), 35–42. https://doi.org/10.1016/j.ridd.2016.07.013

[33] Rebecca Grist, Joanna Porter, and Paul Stallard. 2017. Mental Health Mobile Apps for Preadolescents and Adolescents: A Systematic Review. *Journal of Medical Internet Research* 19, 5 (May 2017), e176. https://doi.org/10.2196/jmir.7332

[34] Monika D Heller, Kurt Roots, Sanjana Srivastava, Jennifer Schumann, Jaideep Srivastava, and T Sigi Hale. 2013. A Machine Learning-Based Analysis of Game Data for Attention Deficit Hyperactivity Disorder Assessment. *Games for Health Journal* 2, 5 (Oct. 2013), 291–298. https://doi.org/10.1089/g4h.2013.0058

[35] Chris Hollis, Caroline J Falconer, Jennifer L Martin, Craig Whittington, Sarah Stockton, Cris Glazebrook, and E Bethan Davies. 2016. Annual Research Review: Digital health interventions for children and young people with mental health problems – a systematic and meta-review. *Journal of Child Psychology and Psychiatry* 58, 4 (Dec. 2016), 474–503. https://doi.org/10.1111/jcpp.12663

[36] Bernd Huber, Richard F. Davis, Allison Cotter, Emily Junkin, Mindy Yard, Stuart Shieber, Elizabeth Brestan-Knight, and Krzysztof Z. Gajos. 2019. SpecialTime: Automatically Detecting Dialogue Acts from Speech to Support Parent-Child Interaction Therapy. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth'19)*. ACM, 139–148. https://doi.org/10.1145/3329189.3329203

[37] James J Hudziak, William Copeland, Catherine Stanger, and Martha Wadsworth. 2004. Screening for DSM-IV externalizing disorders with the Child Behavior Checklist: a receiver-operating characteristic analysis. *Journal of Child Psychology and Psychiatry* 45, 7 (Aug. 2004), 1299–1307. https://doi.org/10.1111/j.1469-7610.2004.00314.x

[38] Netta Iivari and Marianne Kinnula. 2018. Empowering children through design and making: towards protagonist role adoption. In *Proceedings of the 15th Participatory Design Conference: Full Papers - Volume 1 (PDC '18)*. ACM, 1–12. https://doi.org/10.1145/3210586.3210600

[39] QRS International. 2022. *NVivo Qualitative Data Analysis Software*. Retrieved December 20, 2022 from https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

[40] Ole Sejer Iversen, Rachel Charlotte Smith, and Christian Dindler. 2017. Child as Protagonist: Expanding the Role of Children in Participatory Design. In *Proceedings of the 2017 Conference on Interaction Design and Children (IDC '17)*. ACM, 27–37. https://doi.org/10.1145/3078072.3079725

[41] Jacinta Jardine, Caroline Earley, Derek Richards, Ladislav Timulak, Jorge E Palacios, Daniel Duffy, Karen Tierney, and Gavin Doherty. 2020. The Experience of Guided Online Therapy: A Longitudinal, Qualitative Analysis of Client Feedback in a Naturalistic RCT. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. ACM, 1–15. https://doi.org/10.1145/3313831.3376254

[42] Xinlong Jiang, Yiqiang Chen, Wuliang Huang, Teng Zhang, Chenlong Gao, Yunbing Xing, and Yi Zheng. 2020. WeDA: Designing and Evaluating A Scale-driven Wearable Diagnostic Assessment System for Children with ADHD. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. ACM, 1–12. https://doi.org/10.1145/3313831.3376374

[43] Eunkyung Jo, Seora Park, Hyeonseok Bang, Youngeun Hong, Yeni Kim, Jungwon Choi, Bung Nyun Kim, Daniel A Epstein, and Hwajung Hong. 2022. GeniAuti: Toward Data-Driven Interventions to Challenging Behaviors of Autistic Children through Caregivers' Tracking. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (March 2022), 1–27. https://doi.org/10.1145/3512939

[44] Manasa Kalanadhabhatta, Adrelys Mateo Santana, Zhongyang Zhang, Deepak Ganesan, Adam S Grabell, and Tauhidur Rahman. 2022. EarlyScreen: Multiscale Instance Fusion for Predicting Neural Activation and Psychopathology in Preschool Children. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (July 2022), 1–39. https://doi.org/10.1145/3534583

[45] Elizabeth Kaziunas, Mark S Ackerman, Silvia Lindtner, and Joyce M Lee. 2017. Caring through Data: Attending to the Social and Emotional Experiences of Health Datafication. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, 2260–2272. https://doi.org/10.1145/2998181.2998303

[46] Christina Kelley, Bongshin Lee, and Lauren Wilcox. 2017. Self-tracking for Mental Wellness: Understanding Expert Perspectives and Student Experiences. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, 629–641. https://doi.org/10.1145/3025453.3025750

[47] Philip C Kendall and Ronald C Kessler. 2002. The impact of childhood psychopathology interventions on subsequent substance abuse: Policy implications, comments, and recommendations. *Journal of Consulting and Clinical Psychology* 70, 6 (Dec. 2002), 1303–1306. https://doi.org/10.1037/0022-006x.70.6.1303

[48] Julie A. Kientz, Rosa I. Arriaga, and Gregory D. Abowd. 2009. Baby steps: evaluation of a system to support record-keeping for parents of young children. In

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09). ACM, 1713–1722. https://doi.org/10.1145/1518701.1518965

[49] Yoojung Kim, Arpita Bhattacharya, Julie A Kientz, and Jin Ha Lee. 2020. "It Should Be a Game for Fun, Not Exercise": Tensions in Designing Health-Related Features for Pokémon GO. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20).* ACM, 1–13. https://doi.org/10.1145/3313831.3376830

[50] Susanne Kirchner, Dawn K Sakaguchi-Tang, Rebecca Michelson, Sean A Munson, and Julie A Kientz. 2020. "This just felt to me like the right thing to do": Decision-Making Experiences of Parents of Young Children. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference (DIS '20).* ACM, 489–503. https://doi.org/10.1145/3357236.3395466

[51] Robert Kohn, Ali Ahsan Ali, Victor Puac-Polanco, Chantal Figueroa, Victor López-Soto, Kristen Morgan, Sandra Saldivia, and Benjamín Vicente. 2018. Mental health in the Americas: an overview of the treatment gap. *Revista Panamericana de Salud Pública* 42 (2018), e165. https://doi.org/10.26633/rpsp.2018.165

[52] Rachel Kornfield, Jonah Meyerhoff, Hannah Studd, Ananya Bhattacharjee, Joseph Jay Williams, Madhu Reddy, and David C Mohr. 2022. Meeting Users Where They Are: User-centered Design of an Automated Text Messaging Tool to Support the Mental Health of Young Adults. In *CHI Conference on Human Factors in Computing Systems (CHI '22).* ACM, 1–16. https://doi.org/10.1145/3491102.3502046

[53] Deborah A Lupton. 2011. 'The best thing for the baby': Mothers' concepts and experiences related to promoting their infants' health and development. *Health, Risk & Society* 13, 7–8 (Oct. 2011), 637–651. https://doi.org/10.1080/13698575.2011.624179

[54] Regan Lee Mandryk and Max Valentin Birk. 2019. The Potential of Game-Based Digital Biomarkers for Modeling Mental Health. *JMIR Mental Health* 6, 4 (April 2019), e13485. https://doi.org/10.2196/13485

[55] Lyndonna Marrast, David U Himmelstein, and Steffie Woolhandler. 2016. Racial and Ethnic Disparities in Mental Health Care for Children and Young Adults: A National Study. *International Journal of Health Services* 46, 4 (Aug. 2016), 810–824. https://doi.org/10.1177/0020731416662736

[56] Gary Marx and Valerie Steeves. 2010. From the Beginning: Children as Subjects and Agents of Surveillance. *Surveillance & Society* 7, 3/4 (June 2010), 192–230. https://doi.org/10.24908/ss.v7i3/4.4152

[57] Eric J Mash and John Hunsley. 2005. Evidence-Based Assessment of Child and Adolescent Disorders: Issues and Challenges. *Journal of Clinical Child & Adolescent Psychology* 34, 3 (Aug. 2005), 362–379. https://doi.org/10.1207/s15374424jccp3403_1

[58] Kathryn Taaffe McLearn, Jane Knitzer, and Alice S Carter. 2007. *Mental Health: A Neglected Partner in the Healthy Development of Young Children.* American Psychological Association, 233–248. https://doi.org/10.1037/11486-014

[59] Robert R Morris, Kareem Kouddous, Rohan Kshirsagar, and Stephen M Schueller. 2018. Towards an Artificially Empathic Conversational Agent for Mental Health Applications: System Design and User Perceptions. *Journal of Medical Internet Research* 20, 6 (June 2018), e10148. https://doi.org/10.2196/10148

[60] Annika Muehlbradt, Gregory Whiting, Shaun Kane, and Laura Devendorf. 2022. Knitting Access: Exploring Stateful Textiles with People with Disabilities. In *Designing Interactive Systems Conference (DIS '22).* ACM, 1058–1070. https://doi.org/10.1145/3532106.3533551

[61] Lisa K. Mundy, Louise Canterford, Margarita Moreno-Betancur, Monsurul Hoq, Russell M. Viner, Jordana K. Bayer, Petra Lietz, Gerry Redmond, and George C. Patton. 2022. Learning outcomes in primary school children with emotional problems: a prospective cohort study. *Child and Adolescent Mental Health* 28, 3 (Nov. 2022), 377–384. https://doi.org/10.1111/camh.12607

[62] Elizabeth L Murnane, Tara G Walker, Beck Tench, Stephen Voida, and Jaime Snyder. 2018. Personal Informatics in Interpersonal Contexts: Towards the Design of Technology that Supports the Social Ecologies of Long-Term Mental Health Management. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (Nov. 2018), 1–27. https://doi.org/10.1145/3274396

[63] Mario Muñoz-Organero, Lauren Powell, Ben Heller, Val Harpin, and Jack Parker. 2018. Automatic Extraction and Detection of Characteristic Movement Patterns in Children with ADHD Based on a Convolutional Neural Network (CNN) and Acceleration Images. *Sensors* 18, 11 (Nov. 2018), 3924. https://doi.org/10.3390/s18113924

[64] N Nazneen, Agata Rozga, Mario Romero, Addie J Findley, Nathan A Call, Gregory D Abowd, and Rosa I Arriaga. 2011. Supporting parents for in-home capture of problem behaviors of children with developmental disabilities. *Personal and Ubiquitous Computing* 16, 2 (May 2011), 193–207. https://doi.org/10.1007/s00779-011-0385-1

[65] Margaret K Nelson. 2009. *Watching Children: Describing the Use of Baby Monitors on Epinions.com.* Vol. 29. Vanderbilt University Press, 219–238. https://doi.org/10.2307/j.ctv17vf76w.16

[66] A. Noorhidawati, S.Ghazal Ghalebandi, and R. Siti Hajar. 2015. How do young children engage with mobile apps? Cognitive, psychomotor, and affective perspective. *Computers & Education* 87 (Sept. 2015), 385–395. https://doi.org/10.1016/j.compedu.2015.07.005

[67] Jeong Ha Park, Doug Hyun Han, Bung-Nyun Kim, Jae Hoon Cheong, and Young-Sik Lee. 2016. Correlations among Social Anxiety, Self-Esteem, Impulsivity, and Game Genre in Patients with Problematic Online Game Playing. *Psychiatry Investigation* 13, 3 (2016), 297. https://doi.org/10.4306/pi.2016.13.3.297

[68] Soya Park, Jonathan Bragg, Michael Chang, Kevin Larson, and Danielle Bragg. 2022. Exploring Team-Sourced Hyperlinks to Address Navigation Challenges for Low-Vision Readers of Scientific Papers. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (Nov. 2022), 1–23. https://doi.org/10.1145/3555629

[69] Daniel S Pine, Patricia Cohen, Diana Gurley, Judith Brook, and Yuju Ma. 1998. The Risk for Early-Adulthood Anxiety and Depressive Disorders in Adolescents With Anxiety and Depressive Disorders. *Archives of General Psychiatry* 55, 1 (Jan. 1998), 56. https://doi.org/10.1001/archpsyc.55.1.56

[70] Cecil R Reynolds, Randy W Kamphaus, and Kimberly J Vannest. 2015. *BASC3: Behavior assessment system for children.* PscyhCorp Bloomington.

[71] Olivia K Richards, Adrian Choi, and Gabriela Marcu. 2021. Shared Understanding in Care Coordination for Children's Behavioral Health. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (April 2021), 1–25. https://doi.org/10.1145/3449005

[72] Fabio E. G. Santos, Angela P. Z. Bastos, Leila C. V. Andrade, Kate Revoredo, and Paulo Mattos. 2011. Assessment of ADHD through a Computer Game: An Experiment with a Sample of Students. In *2011 Third International Conference on Games and Virtual Worlds for Serious Applications.* IEEE, IEEE, 104–111. https://doi.org/10.1109/vs-games.2011.21

[73] Michael G Sawyer, Leanne Whaites, Joseph M Rey, Philip L Hazell, Brian W Graetz, and Peter Baghurst. 2002. Health-Related Quality of Life of Children and Adolescents With Mental Disorders. *Journal of the American Academy of Child & Adolescent Psychiatry* 41, 5 (May 2002), 530–537. https://doi.org/10.1097/00004583-200205000-00010

[74] Woosuk Seo, Andrew BL Berry, Prachi Bhagane, Sung Won Choi, Ayse G Buyuktur, and Sun Young Park. 2019. Balancing Tensions between Caregiving and Parenting Responsibilities in Pediatric Patient Care. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 1–24. https://doi.org/10.1145/3359255

[75] Ji Youn Shin, Dima Chaar, Catherine Davis, Sung Won Choi, and Hee Rin Lee. 2021. Every Cloud Has a Silver Lining: Exploring Experiential Knowledge and Assets of Family Caregivers. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 1–25. https://doi.org/10.1145/3479560

[76] Hyunjoo Song, Do-Joon Yi, and Hae-Jeong Park. 2020. Validation of a mobile game-based assessment of cognitive control among children and adolescents. *PLOS ONE* 15, 3 (March 2020), e0230498. https://doi.org/10.1371/journal.pone.0230498

[77] Roland Sturm, Jeanne S Ringel, and Tatiana Andreyeva. 2003. Geographic Disparities in Children's Mental Health Care. *Pediatrics* 112, 4 (Oct. 2003), e308–e308. https://doi.org/10.1542/peds.112.4.e308

[78] Hyewon Suh, John R Porter, Alexis Hiniker, and Julie A Kientz. 2014. @BabySteps: design and evaluation of a system for using twitter for tracking children's developmental milestones. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14).* ACM, 2279–2288. https://doi.org/10.1145/2556288.2557386

[79] Hyewon Suh, John R Porter, Robert Racadio, Yi-Chen Sung, and Julie A Kientz. 2016. Baby steps text: Feasibility study of an SMS-based tool for tracking children's developmental progress. In *AMIA Annual Symposium Proceedings*, Vol. 2016. American Medical Informatics Association, 1997.

[80] Kaiwen Sun, Abraham H Mhaidli, Sonakshi Watel, Christopher A Brooks, and Florian Schaub. 2019. It's My Data! Tensions Among Stakeholders of a Learning Analytics Dashboard. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19).* ACM, 1–14. https://doi.org/10.1145/3290605.3300824

[81] Deborah Tatar. 2007. The design tensions framework. *Human–Computer Interaction* 22, 4 (2007), 413–451.

[82] Hiran Thabrew, Theresa Fleming, Sarah Hetrick, and Sally Merry. 2018. Co-design of eHealth Interventions With Children and Young People. *Frontiers in Psychiatry* 9 (Oct. 2018), 481. https://doi.org/10.3389/fpsyt.2018.00481

[83] Janne van Kollenburg, Sander Bogers, Heleen Rutjes, Eva Deckers, Joep Frens, and Caroline Hummels. 2018. Exploring the Value of Parent Tracked Baby Data in Interactions with Healthcare Professionals: A Data-Enabled Design Exploration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18).* ACM. https://doi.org/10.1145/3173574.3173871

[84] Maya Varma, Peter Washington, Brianna Chrisman, Aaron Kline, Emilie Leblanc, Kelley Paskov, Nate Stockham, Jae-Yoon Jung, Min Woo Sun, and Dennis P Wall. 2022. Identification of Social Engagement Indicators Associated With Autism Spectrum Disorder Using a Game-Based Mobile App: Comparative Study of Gaze Fixation and Visual Scanning Methods. *Journal of Medical Internet Research* 24, 2 (Feb. 2022), e31830. https://doi.org/10.2196/31830

[85] Anila Virani, Linda Duffett-Leger, and Nicole Letourneau. 2021. Parents' use of mobile applications in the first year of parenthood: a narrative review of the literature. *Health Technology* 5 (Nov. 2021), 14–14. https://doi.org/10.21037/ht-

20-28

[86] Lauren S Wakschlag, Patrick H Tolan, and Bennett L Leventhal. 2009. Research Review: 'Ain't misbehavin': Towards a developmentally-specified nosology for preschool disruptive behavior. *Journal of Child Psychology and Psychiatry* 51, 1 (Dec. 2009), 3–22. https://doi.org/10.1111/j.1469-7610.2009.02184.x

[87] Jayne Wallace, John McCarthy, Peter C Wright, and Patrick Olivier. 2013. Making design probes work. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3441–3450. https://doi.org/10.1145/2470654.2466473

[88] Junqing Wang, Aisling Ann O'Kane, Nikki Newhouse, Geraint Rhys Sethu-Jones, and Kaya de Barbaro. 2017. Quantified Baby: Parenting and the Use of a Baby Wearable in the Wild. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (Dec. 2017), 1–19. https://doi.org/10.1145/3134743

[89] Xiyue Wang, Kazuki Takashima, Tomoaki Adachi, and Yoshifumi Kitamura. 2021. Can Playing with Toy Blocks Reflect Behavior Problems in Children?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, 1–14. https://doi.org/10.1145/3411764.3445119

[90] Daniel G Whitney and Mark D Peterson. 2019. US National and State-Level Prevalence of Mental Health Disorders and Disparities of Mental Health Care Use in Children. *JAMA Pediatrics* 173, 4 (April 2019), 389. https://doi.org/10.1001/jamapediatrics.2018.5399

[91] Renwen Zhang, Kathryn E. Ringland, Melina Paan, David C. Mohr, and Madhu Reddy. 2021. Designing for Emotional Well-being: Integrating Persuasion and Customization into Mental Health Technologies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, 1–13. https://doi.org/10.1145/3411764.3445771

A DESCRIPTION OF EARLYSCREEN IN THE CLINICIAN SURVEY

Below is a reproduction of the description of EarlyScreen that was presented to mental health practitioners in the survey described in Section 3.1.

> We will now describe **EarlyScreen**, a prototype digital mental health screening tool that has been designed to help parents, caregivers, and clinicians identify specific disorders common in early childhood. EarlyScreen is an additional at-home tool meant to complement existing clinical practices.
>
> One possible use for EarlyScreen is to make the diagnostic intake procedure more efficient and convenient for clinicians and families, particularly low-resource, low-income families. For example, EarlyScreen could be used with a family on the waitlist who urgently needs to pivot to treatment, a low-income family who has difficulty attending multiple in-person intake appointments, or a clinician looking for diagnostic information from a modality other than questionnaires or observation. It could also be used as an additional tool to track changes over the course of therapy.
>
> - *EarlyScreen is a smartphone- or tablet-based "game" that can be played by preschool-aged children that is modeled after existing iPad games.*
> - *The game will induce frustration in children using a clinically-validated paradigm and record facial videos during the process using the tablet's front camera. (In the current iteration, the game involves providing negative feedback on children's choices – see link for more information).*
> - *Facial expressions and head and eye movements are extracted from the captured video and used by machine learning models to predict:*

> - *neural activation within the prefrontal cortex, a region of the brain involved in emotion regulation.*
> - *the child's score on a series of clinically-validated questionnaires to screen for externalizing disorders and ADHD.*
>
> In lab-based tests with 76 participants, a prototype of EarlyScreen could correctly identify 75% of the children exhibiting abnormally low neural activation in the prefrontal cortex during frustration and 77% of children exhibiting normal levels of neural activation. (For context, the Child Behavior Checklist – a well-validated screening tool – correctly identifies 66% of children with exhibiting problematic behavior).
>
> EarlyScreen could also correctly identify 72% of children who scored above clinical thresholds on the CBCL Externalizing disorders, MAP-DB temper loss, and ADHD Inattention and Hyperactivity scales and 76% of children who were below the clinical thresholds.

B SAMPLE CODES MAPPED TO EACH THEME DESCRIBED IN THE FINDINGS

Table 3 lists a few sample codes associated with each theme, grouped by the topics explored in Section 4.

Table 3: Sample Theme-Code Mapping

Topic	Theme	Sample Codes
Concerns and Gaps	Lack of Scaffolding	seeking reassurance, unsure when to seek help
	Limited Resources for Tracking Development	way to measure development, lack of measures
	Incomplete Data	behavioral data, data from other parent
Potential Benefits	Increased Access	in-home use, increased access
	Augmenting Existing Data and Practices	passive information, real-time observation and support
	Frictionless Screening	engaging for children, relevant to existing interests
	Supporting Parents	alleviates pressure, developmental charts
Potential Challenges	Integration & Follow-up	replacing traditional services, ensure follow-up and intervention
	Building Trust	distrust existing technology, concerns about consent
	Minimizing Harm	privacy and data storage, screen time and device use